Contrasting cross-correlation: Meta-analyses of the associations between citations and 13 altmetrics, incorporating moderating variables

Dimity Stephen[1] · Stephan Stahlschmidt[1]

Abstract

Since the Altmetric Manifesto was published in 2010, a large number of studies have quantitatively examined altmetric-citation associations to assess if and how altmetrics differ from citations. These efforts resulted in a diverse array of observations that varies within and between indicators. These studies also differ in the measurement environment, potentially introducing confounding variables. We sought here to clarify this diversity with a meta-analysis and examine if any of several factors moderated the altmetric-citation relationship. We identified relevant literature examining altmetric-citation correlations in the Web of Science and extracted from each study the correlation coefficient, sample size, and potential moderator variables, such as publication year, field, and citation source. We collated 914 correlation coefficients from 111 studies examining 13 altmetric indicators. We fitted multi-level random-effects meta-analysis models for each altmetric indicator individually and examined moderators for Mendeley, Twitter, Altmetric Attention Score, and usage metrics. Mendeley, usage metrics, ResearchGate, Twitter, and peer ratings showed the strongest association with citations, while Facebook, Wikipedia, Google+, blogs, news, and Reddit were only weakly associated with citations. No variable consistently moderated the altmetric-citation relationship, suggesting that, when associations were observed, they were structural and persistent. Altmetrics that did not demonstrate an association evidently differ from citations in the impact construct measured. Hence, our study highlights the diversity among altmetrics channels, as the characteristics of each channel partly define the potential mentions of research in these channels, suggesting a nuanced application of the diverse channels.

✉ Dimity Stephen
  stephen@dzhw.eu

[1] German Centre for Higher Education Research and Science Studies (DZHW), Schützenstraße 6A, Berlin, Germany

Introduction

With the rise of social media and the digitalisation of science, alternative metrics – or altmetrics – have been studied as a new data source for scholarly communication (Sugimoto et al., 2017). Altmetrics are generated from a range of online sources. For example, interactions on social media channels like Twitter and Facebook; mentions of articles in blogs, mainstream news media or Wikipedia; saves of articles to reference managers; or usage metrics such as PDF downloads and page views. Digital Science also introduced the Altmetric Attention Score (AAS), which draws together weighted counts of activity from these and other online sources to generate a composite indicator.

A key feature of altmetrics is their immediacy; altmetrics attention peaks within days to months after publication (Eysenbach, 2011). In comparison, citations – the conventional measure of scientific impact – typically only reach a level that approximates long-term impact after 3 or more years (Wang, 2013). Hence, altmetrics have been discussed as a potential means of impact assessment, either as a substitute for citations due to an observed correlation between both metrics (Costas et al., 2015), or by complementing citations with additional information, such as societal impact (Bornmann, 2014).

The ability of altmetrics to substitute or complement citations has been studied since its suggestion in the Altmetric Manifesto (Priem et al., 2010). To this end, many studies have investigated the correlation between altmetric indicators and citations. However, the results of these studies have been exceedingly diverse. For instance, depending on the sample used, correlations between citations and Twitter mentions have ranged from -0.20 to 0.78 (e.g., Haustein et al., 2014; Malecki, 2015; Xia et al., 2016). We also observe marked variability between altmetric indicators; counts of Mendeley readers and usage metrics are consistently more strongly associated with citations than other sources, such as mentions in blogs and news media (e.g., Amath, 2017; Buttliere & Buder, 2017; Cho, 2021; Gorraiz, Blahous & Wieland 2018). As such, the large number of studies examining the altmetric-citation association has – rather than established a consensus – revealed substantial diversity within and between indicators.

A meta-analytical approach may thus derive clarity from the wealth of studies of altmetric-citation correlations. Three meta-analyses have been conducted in this space to date. Bornmann (2015) examined the pooled correlation between citations and four altmetric data sources. He found weak correlations with Twitter activity (pooled $r = 0.003$, $n = 9$) and mentions in blogs (pooled $r = 0.12$, $n = 9$), and stronger correlations between citations and CiteULike bookmarks (pooled $r = 0.23$, $n = 19$) and Mendeley reader counts (pooled $r = 0.51$, $n = 27$). A meta-analysis by Erdt et al. (2016) spanning nine altmetric sources replicated the result of Bornmann (2015) for blogs ($r = 0.12$, $n = 4$), but identified stronger correlations with Twitter mentions (0.11, $n = 5$) and CiteULike bookmarks (0.29, $n = 13$), and a weaker correlation with Mendeley readers (0.37, $n = 25$). In addition, they observed weak correlations between citations and mentions on Google+ (0.07, $n = 2$), Delicious (0.07, $n = 4$), Wikipedia (0.10, $n = 7$), and Facebook (0.12, $n = 4$), and F1000 ratings (0.23, $n = 25$; Erdt et al., 2016). A third meta-analysis in the health sciences found an overall pooled correlation of 0.19 ($n = 35$) between citations and the AAS (Kolahi et al., 2021). These meta-analyses thus highlight some overarching trends in altmetric-citation associations.

However, a significant number of correlation studies have been published since these meta-analyses were conducted. We thus use the opportunity to examine the altmetric-citation associations using a much larger sample. Additionally, the larger corpus facilitates for the first time a meta-analysis of variables that potentially moderate the relationship

between citations and altmetrics, such as the recency of the study or the discipline examined. As such, we undertake here up-to-date meta-analyses of the association between citations and several altmetric indicators. We also investigate several potential moderators with the aim to crystallise the association irrespective of the variability introduced by these confounders, i.e. controlling for the distinct measurement environments. In undertaking this study, we seek to answer two research questions: (i) what is the strength and direction of the pooled association between several altmetric measures and citations based on the existing literature; and (ii) which characteristics moderate the association between altmetrics and citations, and to what extent? These answers assist in determining whether, and which, altmetrics embody a similar concept of impact as citations, which has implications for how altmetrics are applied in impact assessments.

Method

The study is comprised of two phases: a search and review of the existing literature to identify and extract data from studies that have assessed altmetric-citation correlations, and a meta-analysis of these studies to generate pooled estimates of the association between citations and several altmetric indicators.
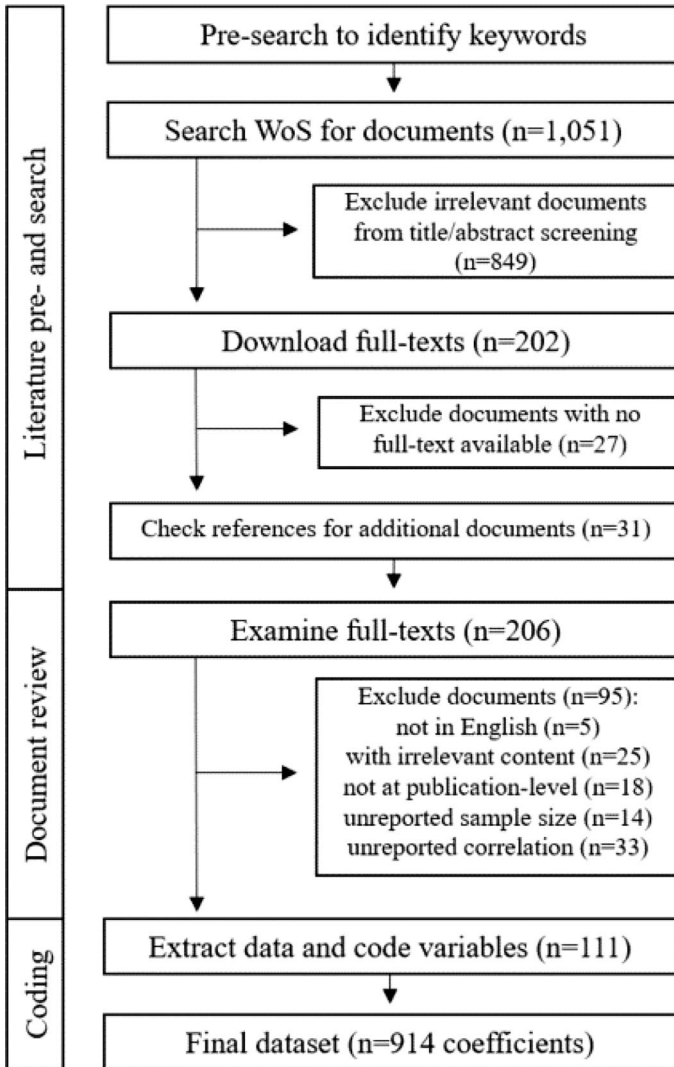
Literature search and review

We carried out the literature search and review over several steps, as shown in Fig. 1. First, we searched the literature about altmetrics and read the titles and abstracts of studies in order to identify the broadest set of keywords that could be used to search for relevant studies. Based on these results, we then searched the Kompetenznetzwerk Bibliometrie's[1] in-house version of the Web of Science (WoS) database for relevant studies published up to April 2022. We searched for articles with abstracts that contained at least one term from each of the following sets: (citation, cite, citing, traditional metric), (relation*, associat*, predict*, correlat*), and (altmetric*, alternative metric*, twitter, facebook, mendeley, tweet, f1000, blog, social media). We included publications of all years and document types. This step identified 1,051 relevant documents. However, based on a review of their titles and/or abstracts, we excluded 849 documents as they were irrelevant to our study.

We downloaded the full-texts of the remaining 202 documents, excluding 27 documents that we could not access. We identified an additional 31 relevant documents from the studies' reference lists and incorporated these. We then reviewed the full-texts of these 206 documents against the following inclusion criteria: the study was empirical; examined the association between at least one altmetric indicator and citations; examined this association at the document level, i.e. not author, journal, etc.; reported a Pearson, Spearman or Kendall correlation or R^2 statistic; reported the sample size; and was written in English.

We then extracted from each of the 111 remaining studies the correlation coefficient, the sample size, and the altmetric data source used. Each study typically reported more than one correlation coefficient as they usually contained multiple samples based on publications from different publication years, disciplines, or citation sources and reported a coefficient for each sample. We did not differentiate between Pearson and Spearman correlations

---

[1] https://bibliometrie.info/

Fig. 1 Flowchart of the steps taken for the literature search, review, and coding

as the statistics are comparable (Shen et al., 2021). We converted Kendall's correlation coefficients to Spearman's *r* based on the conversion table provided by Gilpin (1993). One author (DS) also coded the following additional variables in each study to examine them as moderators of the altmetric-citation association: (i) the study's publication year, (ii) the publication years of articles in the study's sample, (iii) the citation data source used, (iv) the discipline of the sampled publications concorded to the OECD's Fields of Science and Technology (FOS), (v) a binary indicator of whether the sample included only articles that had non-zero altmetric values or citations, (vi) a binary indicator of whether the sample included only articles that were highly cited or had high altmetric values, and ordinal variables of the time between publication and when the vii) altmetrics data and viii) citation data were collected (0–1 years, 2–4 years, 5+years).

We coded the altmetric data sources to AAS, blogs, Facebook, Google+, Mendeley, other reference managers (i.e., CiteULike, Connotea, and Delicious), news outlets, peer rating (e.g., F1000), Reddit, ResearchGate, Twitter, usage metrics (e.g., views or downloads from journal websites), Wikipedia, and "other". "Other" encompassed data from comments on PLoS webpages, bookmarks on CN3, and mentions on forums, Q&A, and LinkedIn pages. We classified the citation data sources as WoS, Scopus, or "other". This latter category combined observations from the Chinese National Knowledge Infrastructure, Citebase, CiteSeer, Crossref, Dimensions, Google Scholar, iCite, PubMed, PMC Europe, and journal websites. We aligned the disciplines used in each study from the native WoS and Scopus classifications to the FOS classification based on concordances provided by Clarivate Analytics and Elsevier, respectively. When studies used other discipline classifications, we manually assigned these samples to the FOS. The FOS classification consists of *Agricultural sciences, Medical and health sciences, Natural sciences, Engineering and technology, Social sciences,* and *Humanities* fields. Samples that did not examine a specific field but studied a general sample of articles were classified to an *All fields* category. We determined the altmetric and citation windows based on the publication year of the sample and the year in which the altmetrics or citation data were collected, where this information was reported in the study. For samples where a range of publication and or collection years was used, we allocated this sample to the later applicable period, e.g., a 3–6-year window was allocated to the 5 + years category.

Meta-analytic method

The second phase of the study consisted of conducting meta-analyses to examine the strength of the pooled association between citations and the individual altmetric indicators. As most studies contributed more than one observation, the samples were often not independent. To account for this lack of independence, we conducted multi-level meta-analyses. While meta-analyses inherently assume a nested structure of the observations within studies and thus two levels of variance (random sampling error and between-study heterogeneity), multi-level models contain a third level which assumes that the observations are correlated due to their clustering within studies. If not accounted for, this lack of independence in the sample may be interpreted as less heterogeneity and a false-positive outcome (Harrer et al., 2021).

The nature of correlations as restricted to a range of − 1 to 1 can introduce bias in estimates of the standard error due to a compression of the value range (Harrer et al., 2021). Consequently, we first transformed the observed correlations to Fisher's z and also calculated the standard error of the observation from this statistic. We then used the metafor R package (Viechtbauer, 2010) to fit multi-level random-effects meta-analysis models and estimated variance using restricted maximum likelihood procedures. We fit one model per altmetrics indicator. For indicators with more than 50 observations (Mendeley, AAS, usage metrics, and Twitter), we also included the aforementioned study characteristics to assess whether they moderated the altmetric-citations relationship.

Before fitting the models, we tested the correlation between the two continuous year-related variables (study's publication year and years studied in sample) and the Spearman rank order correlation (base R package stats) between the two ordinal citation and altmetric window variables to avoid over-fitting the model due to multi-collinearity between moderators. For all altmetric indicators, the study's publication year and years sampled were moderately to strongly correlated ($r = 0.52$–0.89), as were the citation and altmetric windows
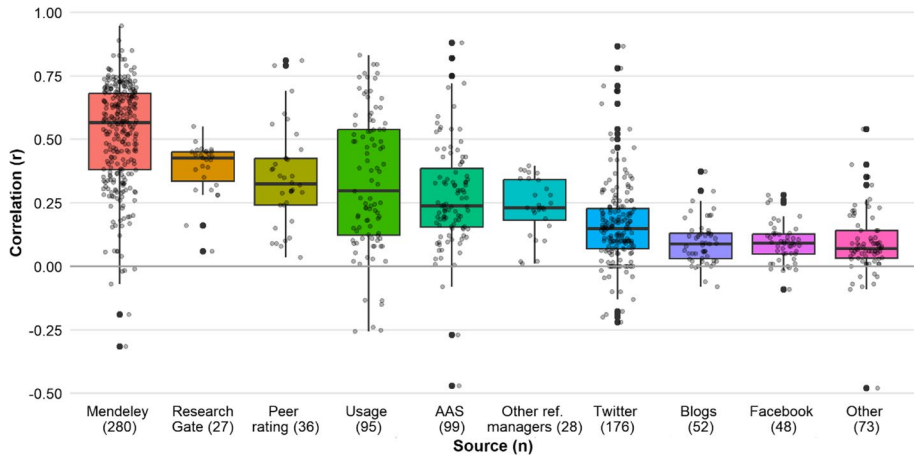
Fig. 2 Distribution of observed altmetric-citation correlations by altmetric source

used ($r = 0.40$–1.00). We thus examined only the sampled year and the citation window as the sampled year better captured the time period analysed, and the citation window had slightly fewer missing data. We centred the sampled year variable before including it in the analysis to improve the interpretation of the intercept. We also excluded the binary indicator of whether the sample included only articles with high altmetrics or citations as only a small number of observations focused on high-value samples. As such, the final set of moderators examined were the (i) sampled year, (ii) OECD field, (iii) citation data source, (iv) whether the sample included only articles with non-zero altmetric or citation values, and (v) the citation window.

Following the method recommended by Assink and Wibbelink (2016), we fit a model for each moderator individually to assess its effect on the association. We then fit a final model for each indicator incorporating the moderators found to be significantly influential. As the *All fields* category was non-specific and could thus impede interpretation of the results, we excluded these observations from models that found field was a significant moderator. For indicators with fewer than 50 observations, we fit only a multi-level model without any moderators. We converted the Fisher's *z* values back to correlation coefficients (*r*) when reporting results, for ease of interpretation. As we used multi-level meta-analyses, we did not test for publication bias because the standard tests, such as Egger's regression and Trim and Fill, have been found to have limited ability to detect publication bias in such cases and no suitable alternative is yet available (Assink & Wibbelink, 2016; Rodgers & Pustejovsky, 2021).

Results

The final number of studies included in our sample was 111 and the total number of correlation coefficients extracted was 914. A list of the studies included can be found in Table S1 of the Supplementary Material. The number of publications included in the samples ranged from 3 to 3,808,747, with a mean of 41,665. We excluded from analyses observations from the 3 studies with sample sizes of 3. The Spearman correlations observed ranged from
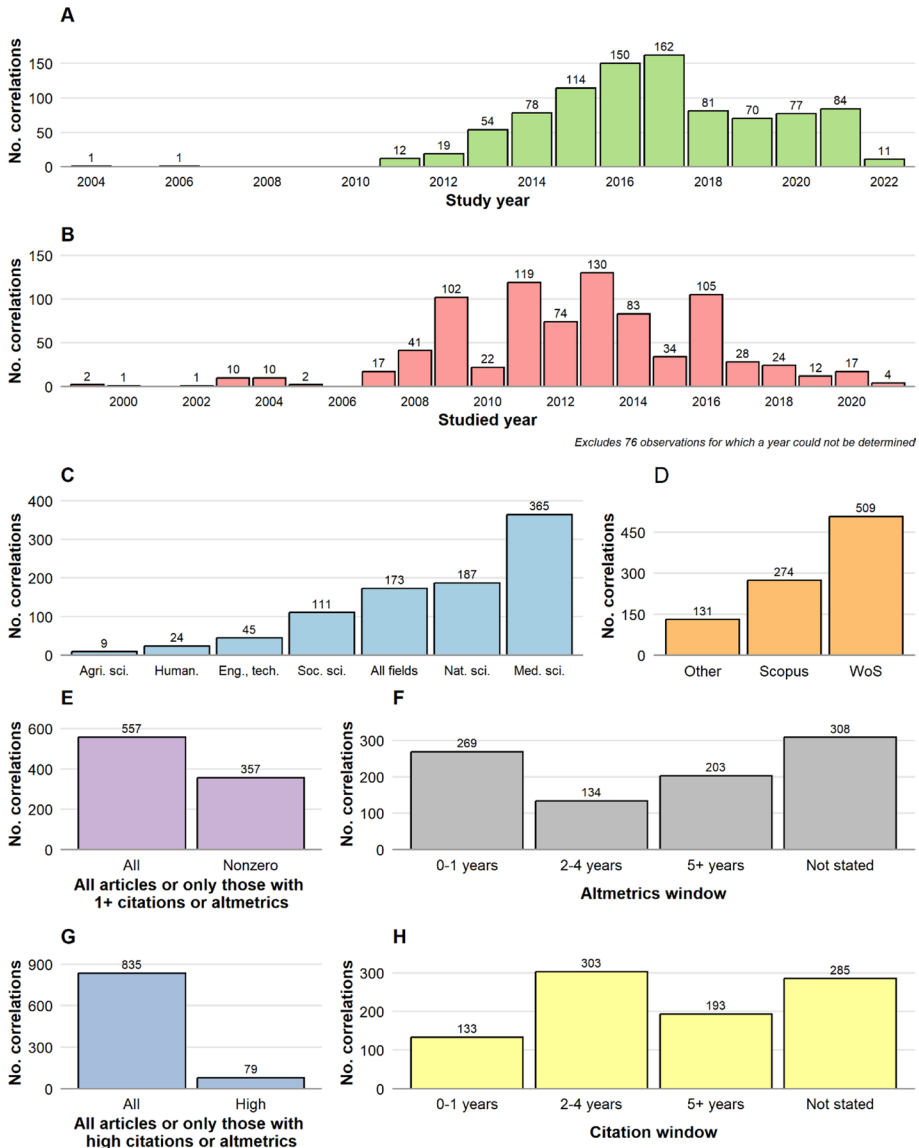
Fig. 3 The number of correlation coefficients extracted from the sample by study characteristics

−0.48 to 0.95, with a mean of 0.30. However, as shown in Fig. 2, the number of observations and the distributions of the correlations varied substantially by the source of altmetric data used.

An overview of the number of observations with particular study characteristics is shown in Fig. 3. Here we can see that, over time, the number of correlation studies conducted steadily rose until reaching 162 in 2017 then halved to approximately that number in the years since (panel A). The years studied were more sporadic, but tended to be concentrated between 2009 and 2016 (panel B). In terms of fields (panel C),

associations have been studied most amongst *Medical and health science* articles (365, 39.9%) and using WoS as the citation data source (509, 55.7%; panel D). Studies tended to examine the altmetrics of articles within a year of publication and citations within 2–4 years (panels F and H), corresponding with the common standard to apply a 3-year citation window and the interest in the immediacy of altmetrics. However, many studies did not report the citation or altmetric windows or dates of data collection required for us to calculate them. Around a third of observations (357, 39.1%) included only articles that had at least 1 citation or altmetric activity, while nearly two-thirds of observations (557, 60.9%; panel E) did not use this restriction. Finally, as mentioned, only a small number of observations (79, 8.6%; panel G) included only articles with high citations or altmetrics, hence we removed it as a potential moderator variable.

We first examined the pooled association between Twitter and citations based on 176 observations of the Twitter-citations correlation collected from 31 studies. The distribution of observations by potential moderator variables is shown in Fig. 4. In modelling the moderators individually, we identified that citation source and citation window were significant moderators and the final multi-level model demonstrated a significant moderating effect ($F(4, 143) = 3.21$, $p < 0.05$). The final model excluded 26 observations from 9 studies that were missing data for the citation window and 2 observations with sample sizes of 3.

Compared to a reference group of "other" citation sources, including Crossref, Dimensions, PubMed and Google Scholar, both Scopus and WoS were associated with weaker Twitter-citation associations, but only WoS was a significant modifier ($r = -0.24$, 95% CIs $= -0.39$, -0.07; $p < 0.05$). The citation window used did not modify the relationship in the final model. The overall pooled correlation estimate was $r = 0.36$ (95% CIs $= 0.19$, 0.50; $p < 0.00$). The majority of variance was attributable to heterogeneity within studies ($I^2 Level2 = 92.7\%$), while only 7.1% of variance stemmed from between-study differences. This may occur as four studies contributed large numbers of observations (i.e. 14–61 observations) for varying samples.

Fifty studies contributed 99 observations of the AAS-citations association. However, 20 observations were excluded due to missing sample year or non-specific field information. The observed associations disaggregated by study characteristics are shown in Fig. 5. The years sampled and field significantly moderated the AAS-citation relationship ($F(3, 75) = 3.02$, $p < 0.05$). No study tested correlations for the agricultural sciences or humanities, so we examined here only 3 fields, using social sciences as the reference group. Natural sciences publications were associated with significantly stronger correlations between AAS and citations ($r = 0.43$, 95% CIs $= 0.07$, 0.69; $p < 0.05$). The sample year was not a significant moderator in this model. Overall, the pooled correlation was $r = 0.18$ (95% CIs $= -0.01$, 0.36; $p = 0.07$) and was not significant. The variance within-studies was also higher ($I^2 Level2 = 71.3\%$) than the variance between-studies ($I^2 Level3 = 24.9\%$). The estimates, standard errors, and significance values for the moderators for the Twitter and AAS models are shown in Table 1.

Thirty-seven studies provided 279 observations of Mendeley-citation correlations. None of the potential moderators significantly influenced this relationship. As such, we fit a three-level random effects model without moderators. The pooled correlation was $r = 0.54$ (95% CIs $= 0.49$, 0.58; $p < 0.001$). Nearly 40% of the variance was attributable to heterogeneity within studies ($I^2 Level2 = 38.5\%$), while between-study differences accounted for nearly two-thirds of variability ($I^2 Level3 = 61.4\%$). The associations observed for Mendeley, disaggregated by study characteristics, are shown in Fig. S1 in the Supplementary Material.
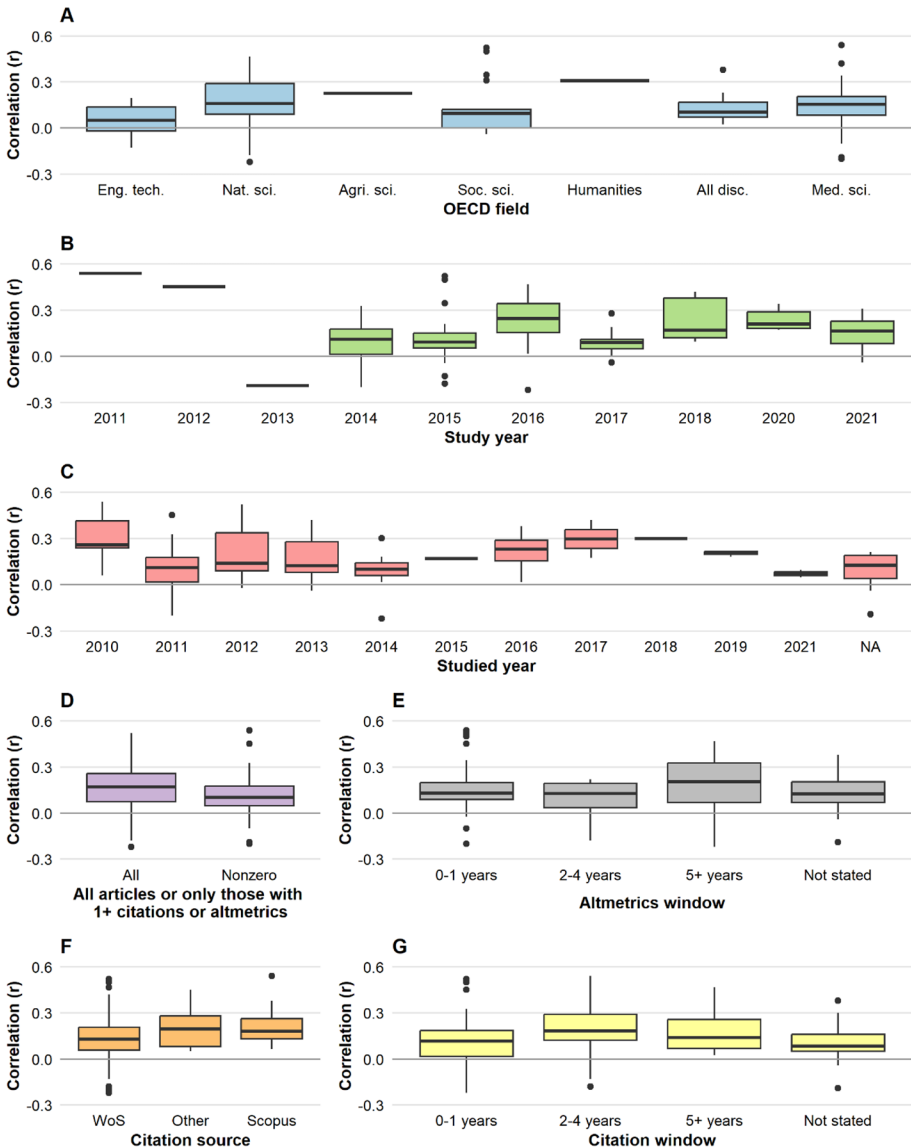
Fig. 4 Distributions of correlations between citations and Twitter activity with potential moderator variables ($n = 176$)

We retrieved 95 observations of the usage metrics-citations correlation from 19 studies. No moderators significantly influenced the relationship and so we fit a multi-level model without moderators. The overall pooled correlation estimate was $r = 0.45$ (95% CIs $= 0.32, 0.55$; $p < 0.01$) and nearly two-thirds of variance was attributable to between-study ($I^2 Level3 = 61.6\%$) and one-third to within-study heterogeneity ($I^2 Level2 = 38.4\%$). The associations observed for usage metrics, disaggregated by study characteristics, are shown in Fig. S2 in the Supplementary Material.
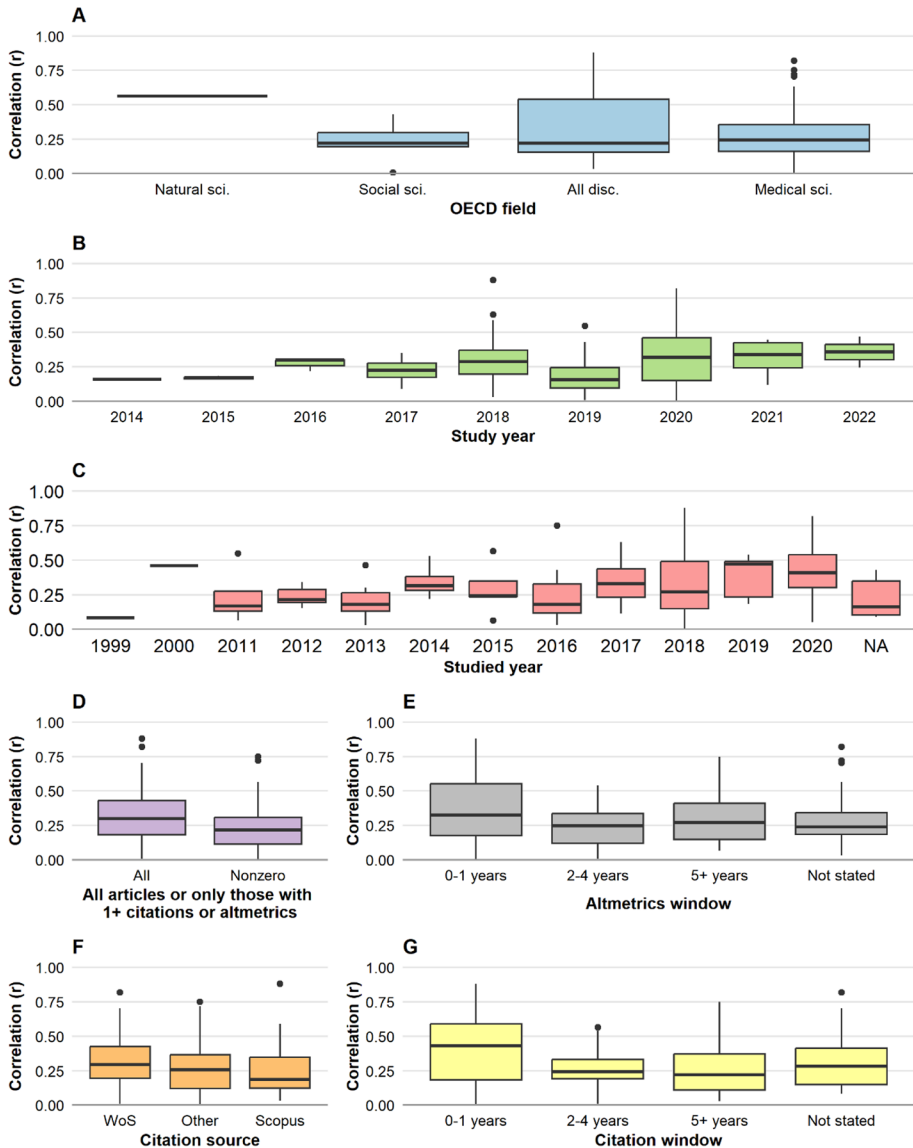
Fig. 5 Distributions of correlations between citations and AAS activity with potential moderator variables (n=99)

For the remaining altmetrics indicators – ResearchGate, peer ratings, blogs, other reference managers, news, Facebook, Wikipedia, Google +, and Reddit – that had fewer than 50 observations, we fit multi-level models without moderators. Figure 6 shows for all altmetric indicators the number of observations, the pooled correlation estimate, the associated confidence intervals and significance indicator, and the percentage of variance attributable to each level of the models. Here we see that there was a statistically significant pooled association with citations for all altmetric indicators except the AAS and mentions

Table 1 Pooled estimates, confidence intervals, and significance values for meta-analysis models with moderators

Altmetric		Estimate(r)	Standard error	Lower 95% CI	Upper 95% CI	p
AAS	Intercept	0.18	0.10	− 0.01	0.36	0.07
AAS	Sampled year (centred)	0.01	0.01	0.00	0.03	0.07
AAS	Field: Medical sciences	0.13	0.10	− 0.08	0.32	0.21
AAS	Field: Natural sciences	0.43	0.19	0.07	0.69	0.02*
Twitter	Intercept	0.36	0.09	0.19	0.50	0.00***
Twitter	Citations: Scopus	− 0.20	0.10	− 0.38	0.00	0.05
Twitter	Citations: WoS	− 0.24	0.09	− 0.39	− 0.07	0.01*
Twitter	Cit. window: 2-4 yr	0.07	0.04	− 0.01	0.14	0.08
Twitter	Cit. window: 5 yr+	0.04	0.05	− 0.06	0.15	0.42

*** $=p<0.001$, ** $=p<0.01$, * $=p<0.05$
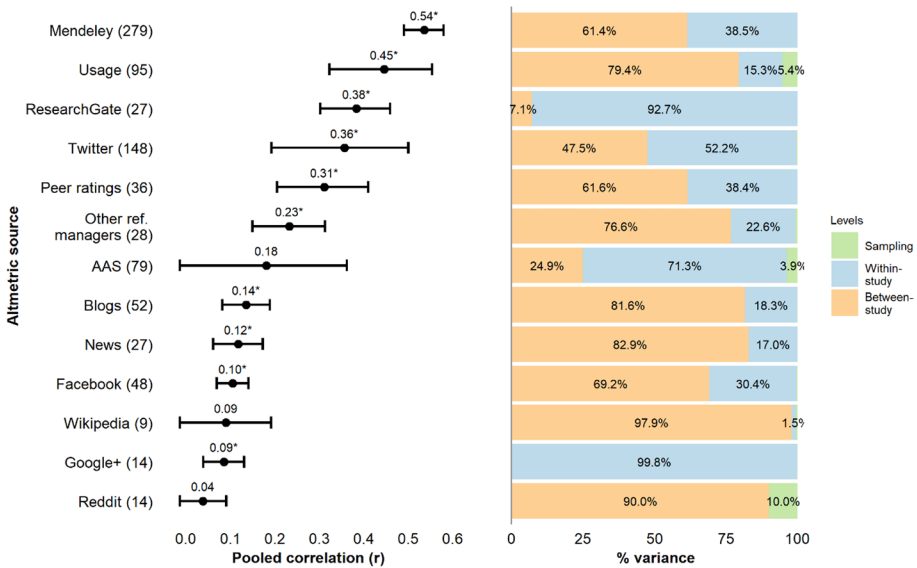
Fig. 6 Number of observations, pooled correlation estimates, 95% CIs, and percentage of variance attributable to model levels for all altmetric indicators. * $=p<0.01$

on Wikipedia and Reddit. However, statistical significance does not necessarily translate to a substantial effect and the strength of these associations varied greatly. The modelled correlation for Mendeley, usage statistics, ResearchGate, Twitter, and peer ratings was much stronger than the modelled correlations for Google +, Facebook, news, and blog mentions.

Comparing the modelled correlation estimates in Fig. 6 with the simple empirical observations of Fig. 2, the clarifying effect of the meta-analysis – which accounts for random sampling errors, between-study heterogeneity, and a lack of independence between observations from the same study – becomes apparent. The pooled Mendeley correlation shows the same high association with citations observed in Fig. 2, but with substantially

reduced variance/uncertainty surrounding the pronounced similarity between both metrics. Usage metrics show a substantially elevated pooled correlation in the modelled values of Fig. 6 compared to the observations in Fig. 2. Accounting for the nested data structure reveals a much higher average association, which is accompanied by sizeable variance. Still, usage metrics have a higher-than-expected resemblance to citations in the model. The same observation holds for Twitter, where the unchecked effects of sampling error, study heterogeneity and lack of independence mask a relatively high pooled correlation of 0.35, although again accompanied by sizeable uncertainty. On the contrary, the meta-analysis results show a reduced average association for the AAS, along with the largest variance among all channels, rendering the correlation statistically insignificant. The composite nature of this indicator building upon individual altmetric channels with varying associations to citations may cause this high variance and resulting uncertainty in its relation to citations, and its own informational value.

Discussion

In this study we conducted meta-analyses to clarify the strength and direction of associations between citations and several altmetric indicators. In an extension to previous studies, we also examined the moderating effect of a number of common study characteristics. Overall, we observed significant positive associations between citations and most altmetric indicators. These associations were substantial for activity on Mendeley, ResearchGate, Twitter, and usage metrics and peer ratings. In contrast, there were only weak (although statistically significant) associations between citations and mentions in blogs, news, Facebook, Google+, and other reference managers, and no significant association with AAS, Wikipedia, or Reddit activity. These results align with findings from previous meta-analyses for Mendeley, blogs, Facebook, Wikipedia, Google+, and AAS (Bornmann, 2015; Erdt et al., 2016; Kolahi et al., 2021). However, we observed a stronger association with Twitter (0.36) than was previously reported (0.003–0.11; Bornmann, 2015; Erdt et al., 2016). As such, altmetric indicators generally demonstrate associations with citations, but to varying degrees depending on the specific altmetric channel.

The variability in the strength of altmetrics' associations with citations highlights that altmetrics do not necessarily encapsulate the same form of impact as each other or as citations. For instance, using factor analysis of variables measuring impact, Bornmann and Haunschild (2018) found that Mendeley readership and citations loaded onto a single factor that was significantly associated with research quality, while Twitter activity loaded onto a separate factor unrelated to research quality. Similarly, Wooldridge and King (2019) determined that peer ratings of a department's research quality in the UK's Research Excellence Framework exercise were significantly related to the department's citation rate, while ratings of societal impact more closely aligned with altmetric attention. As such, there appears to be at least two kinds of impact captured by altmetrics: the reach of a paper in channels that similarly value research quality and are generally used by academic audiences, and the reach of a paper into the more general audiences of social media outlets. We perhaps see this reflected here in the lack of significant association and broad confidence intervals for the AAS, stemming from its combination of data from multiple types of sources. Altmetrics channels that have a large overlap in their user base with an academic audience, e.g., usage metrics, Mendeley, ResearchGate, thus

tend to have the strongest associations with citations, while altmetric channels with a low association seem to be governed by communication patterns outside of purely academic correspondence.

In terms of moderator variables, interestingly, we observed very limited effects of characteristics on the altmetric-citation relationships. The sampled year, field, citation data source, use of articles with non-zero altmetric or citation values, and the citation window had no effect on the association of Mendeley activity or usage metrics with citations. The citation source used moderated the association between Twitter activity and citations, with weaker associations observed between tweets and citations from WoS compared to other citations sources. This effect may arise as the base research-oriented coverage of WoS aligns less with the interests of the mixed general and academic audience of Twitter than data sources with broader coverage including also applied research, such as Dimensions (Stahlschmidt & Stephen, 2022). Field was relevant for AAS, with stronger, positive associations identified for natural science publications compared to social science publications. However, overall, the limited effects of potentially moderating variables suggests that the observed relationships between altmetrics and citations are structural and persistent over time, fields, and citation sources.

There are a few limitations in our study to note. First, as we drew our observations from independent studies, there is some variance in how altmetrics and citations were measured. For instance, usage metrics were calculated as page views, PDF downloads, or a combination thereof, while Twitter activity may have included tweets, retweets, saves, or a combination thereof. Similarly, occasionally citations were calculated as the annual average over a time period, rather than the number within a time period. These differences may have influenced the strength of associations we observed. Secondly, publication bias – the tendency for non-significant findings to not make their way into journals – could be present, leading us to over-estimate the associations in our meta-analyses. Unfortunately, the existing methods to assess the likelihood of publication bias are not recommended for use with multi-level meta-analysis models. However, we anticipate a limited effect of publication bias in our sample as the studies typically examined multiple samples across, e.g., disciplines, citations sources, altmetrics, and reported all the outcomes, rather than just those that were statistically significant. We noted that 18.4% (168) of observations in our sample were not statistically significant, 30.3% (277) did not have their significance reported, and 51.3% (469) observations were statistically significant.

In conclusion, our meta-analyses identified positive relationships between several altmetrics indicators and citations. These modelled associations were strongest for Mendeley, usage metrics, ResearchGate and Twitter activity, likely due to the large overlap in the users of these platforms and academic actors who cite research. Compared to the simple empirical observations shown in Fig. 2, our meta-analytical approach shifted the ordering of altmetric indicators in terms of strength of association with citations, after controlling for random sampling, study heterogeneity, and dependent observations. Further, the associations we observed were largely persistent over time, field, and citation source. As such, the study was able to make use of the large corpus developed by scholars in the field to identify the overall association of the diverse altmetrics channels with citations, allowing future work to apply the established associations, and motivate further analysis on the communication structure of altmetric channels that diverge from citations.

Declarations

Competing interests The authors have no relevant financial or non-financial interests to disclose.

References

Amath, A., Ambacher, K., Leddy, J. J., Wood, T. J., & Ramnana, C. J. (2017). Comparing alternative and traditional dissemination metrics in medical education. *Medical Education, 51*, 935–941. https://doi.org/10.1111/medu.13359

Assink, M., & Wibbelink, C. (2016). Fitting three-level meta-analytic models in R: A step-by-step tutorial. *The Quantitative Methods for Psychology, 12*, 154–174. https://doi.org/10.20982/tqmp.12.3.p154

Bornmann, L. (2014). Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of altmetrics. *Journal of Informetrics, 8*, 895–903. https://doi.org/10.1016/j.joi.2014.09.005

Bornmann, L. (2015). Alternative metrics in scientometrics: A meta-analysis of research into three altmetrics. *Scientometrics, 103*, 1123–1144. https://doi.org/10.1007/s11192-015-1565-y

Bornmann, L., & Haunschild, R. (2018). Do altmetrics correlate with the quality of papers? A large-scale empirical study based on F1000 Prime data. *PLoS ONE, 13*(5), e0197133. https://doi.org/10.1371/journal.pone.0197133

Buttliere, B., & Buder, J. (2017). Personalizing papers using altmetrics: Comparing paper 'quality' or 'impact' to person 'intelligence' or 'personality.' *Scientometrics, 111*, 219–239. https://doi.org/10.1007/s11192-017-2246-9

Cho, J. (2021). Altmetrics analysis of highly cited academic papers in the field of library and information science. *Scientometrics, 126*, 7623–7635. https://doi.org/10.1007/s11192-021-04084-w

Costas, R., Zahedi, Z., & Wouters, P. (2015). Do "altmetrics" correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. *Journal of the Association for Information Science and Technology, 66*, 2003–2019. https://doi.org/10.1002/asi.23309

Erdt, M., Nagarajan, A., Sin, S.-C.J., & Theng, Y.-L. (2016). Altmetrics: An analysis of the state-of-the-art in measuring research impact on social media. *Scientometrics, 109*, 1117–1166. https://doi.org/10.1007/s11192-016-2077-0

Eysenbach, G. (2011). Can tweets predict citations? Metrics of social impact based on Twitter and correlation with traditional metrics of scientific impact. *Journal of Medical Internet Research, 13*, e123. https://doi.org/10.2196/jmir.2012

Gilpin, A. R. (1993). Table for conversion of Kendall's tau to Spearman's rho within the context of measures of magnitude of effect for meta-analysis. *Educational and Psychological Measurement, 53*, 87–92. https://doi.org/10.1177/0013164493053001007

Gorraiz, J., Blahous, B., & Wieland, M. (2018). Monitoring the broader impact of the journal publication output on country level: A case study for Austria. In Erdt M., Sesagiri Raamkumar A., Rasmussen E., & YL. Theng (Eds.), *Altmetrics for research outputs measurement and scholarly information management* (pp. 39–62). Springer. https://doi.org/10.1007/978-981-13-1053-9_4

Harrer, M., Cuijpers, P., Furukawa, T., & Ebert, D. (2021). *Doing meta-analysis with R: A hands-on guide* (1st ed.). Chapman & Hall/CRC Press.

Haustein, S., Peters, I., Sugimoto, C., 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*, 656–669. https://doi.org/10.1002/asi.23101

Kolahi, J., Khazaei, S., Iranmanesh, P., Kim, J., Bang, H., & Khademi, A. (2021). Meta-analysis of correlations between altmetric attention score and citations in health sciences. *BioMed Research International*. https://doi.org/10.1155/2021/6680764

Malecki, A. (2015). *PubMed and arXiv vs. Gold Open Access: Citation, Mendeley, and Twitter uptake of academic articles of Iran* (Salah A. A., Tonta Y., Akdag Salah A. A., C. Sugimoto, & U. Al, Eds.; pp. 46–58). Bogaziçi University Printhouse.

J. Priem, D. Taraborelli, P. Groth, C. Neylon (2010), *Altmetrics: A manifesto*, 26 October 2010. http://altmetrics.org/manifesto

Rodgers, M., & Pustejovsky, J. E. (2021). Evaluating meta-analytic methods to detect selective reporting in the presence of dependent effect sizes. *Psychological Methods, 26*, 141–160. https://doi.org/10.1037/met0000300

Shen, H., Xie, J., Li, J., & Cheng, Y. (2021). The correlation between scientific collaboration and citation count at the paper level: A meta-analysis. *Scientometrics, 126*, 3443–3470. https://doi.org/10.1007/s11192-021-03888-0

Stahlschmidt, S., & Stephen, D. (2022). From indexation policies through citation networks to normalized citation impacts: Web of Science, Scopus, and Dimensions as varying resonance chambers. *Scientometrics, 127*, 2413–2431. https://doi.org/10.1007/s11192-022-04309-6

Sugimoto, C., Work, S., Larivière, V., & Haustein, S. (2017). Scholarly use of social media and altmetrics: A review of the literature. *Journal of the Association for Information Science and Technology, 68*, 2037–2062. https://doi.org/10.1002/asi.23833

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Wang, J. (2013). Citation time window choice for research impact evaluation. *Scientometrics, 94*(3), 851–872. https://doi.org/10.1007/s11192-012-0775-9

Wooldridge, J., & King, M. B. (2019). Altmetric scores: An early indicator of research impact. *Journal of the Association for Information Science and Technology, 70*(3), 271–282. https://doi.org/10.1002/asi.24122

Xia, F., Su, X., Wang, W., Zhang, C., Ning, Z., & Lee, I. (2016). Bibliographic analysis of nature based on Twitter and Facebook altmetrics data. *PLoS ONE, 11*, e0165997. https://doi.org/10.1371/journal.pone.0165997