Reliability and accuracy of altmetric providers: a comparison among Altmetric.com, PlumX and Crossref Event Data

José Luis Ortega

Institute for Advanced Social Studies (IESA-CSIC), Córdoba, Spain, jortega@orgc.csic.es

Abstract

The aim of this study is to measure differences between three relevant altmetric providers: Altmetric.com, PlumX and Crossref Event Data (CED). More than 67,000 research papers, initially extracted from PlumX, were searched in Altmetric.com and CED to compare their counts. Differences between services were analyzed regarding the number of documents with an altmetric event and the counting differences in each metric. Results show that Altmetric.com is the provider with the best coverage of blog posts, news and tweets; while PlumX better collects Mendeley readers; and CED is the site that extracts more Wikipedia citations. The study concludes that there are important counting differences due to technical errors and extracting criteria. The article recommends the use of specific services for the analysis of particular metrics. While, it should be mandatory to employ the combination of several providers, if we want to perform an overall analysis.

Keywords: Altmetric.com, PlumX, Crossref Event Data, Altmetrics, data providers

Introduction

Since the publication of the Altmetrics Manifesto (Priem et al., 2010), a huge number of academic publications have been released around the world on alternative metrics (Holmberg, 2015; Tattersall, 2016; Roemer and Borchardt, 2015). The main objective of these altmetric studies has been to understand the meaning of these metrics and their possible application to research evaluation (Erdt et al., 2016; Sugimoto et al., 2017). In this form, a first step has been to compare and put into context these metrics with regard to bibliometric indicators (Thelwall et al., 2013; Costas et al., 2015). The poor correlations found lead us to think, for now, that these metrics express a new dimension of the academic impact, which is closer to the popularization of science, societal impact and research dissemination (Ortega, 2016).

Altmetric studies have been supported by data produced by a wide scope of services that track the use, mention, sharing and citation of research papers in social networks, publisher platforms, repositories, etc. Faced with this great amount of metrics and sources, altmetric providers have been a key element in the consolidation of this young research field, gathering alternative metrics on a single platform. These databases have made easier the comparison and study of different metrics, putting them in relation to each other and to the citation impact. Based on web syndication and open data, events on the Web about academic outputs are collected by these aggregators using several identifiers (DOIs, ISBN, URLs, etc.). The

counting of these events introduces a new perspective about the social and online impact of these academic results. Today, these services are inserted in the scholarly communication system providing altmetric information to journals (PLOS, Nature), publisher platforms (ScienceDirect, SpringerLink) and digital libraries.

However, and in spite of the importance of these instruments for altmetric studies, there are a few studies that have investigated the consistency and reliability of altmetrics across data providers (Jobmann et al., 2014; Zahedi et al., 2015). This is an important gap because many of the current results might be limited to the data supplied by these aggregators. The aim of this study is to fill this gap and to analyze the coverage and counting differences between three representative data providers: Altmetric.com, PlumX and CED.

Related Research

Altmetric providers constitute an indispensable part in the altmetric research because they supply data for these studies. Some works have described and explored the functionalities of these services, as a way to explain their advantages and drawbacks. At this point, it is worth mentioning the studies of Champieux (2015) and Lindsay (2016) about PlumX, who make a descriptive analysis of its functionalities; the analyses performed by Adie and Roe (2013) and Trueger et al. (2015) about Altmetric.com and the serious criticisms of Gumpenberger et al. (2016) about the Altmetric Score. There are no studies to date about Crossref Event Data (CED) because this service is still in beta and it has not been officially released yet. However, other studies have analyzed the coverage of these services and have described the proportion of altmetric events in several samples. Robinson-García et al. (2014) analyzed the coverage of Altmetric.com and they found that 87.1% of articles had at least one tweet and 64.8% one Mendeley reader. In a similar way, Bornmann (2014) explored a set of articles from Altmetric.com and he observed that 71% of articles were tweeted and a moderated proportion of documents were mentioned in blogs (16%) and news (13%). Fraumann et al. (2015) also described the distribution of blogs and news covered by Altmetric.com and they found an important bias towards U. S. sites. According to PlumX, there are no studies that have treated its coverage. We can only mention the recent work of Torres-Salinas et al. (2017) about the collection of books and the assessment of these materials, and the study of Ortega (2016) about the tweeting of research articles using PlumX data.

However, much fewer papers have carried out comparative studies between data providers. Jobmann et al. (2014) were the first ones to compare the altmetric coverage and counts of ImpactStory, Altmetric Explorer, Plum Analytics and Webometric Analyst by research areas. Their results show important divergences between services. Plum Analytics is the platform that better covers Mendeley and Facebook data, while Altmetric.com stands out gathering blogs, news and CiteULike data. Zahedi et al. (2015) explored the consistency of data across three altmetric aggregators: Altmetric.com, Mendeley and Lagotto. They also detected significant differences, finding that Altmetric.com gathers more tweets, but it is less accurate collecting Mendeley readers. Baessa et al. (2015) evaluated several altmetric providers for their institutional repository and they recognized that Altmetric.com has a better coverage of blogs, news and government documents, while PlumX is most exhaustive covering different formats. Kraker et al. (2015) studied the collection of research data in three data sources: Figshare,

PlumX and ImpactStory. They observed that PlumX detects considerably more items in social media and also finds higher altmetric scores than ImpactStory. Peters et al. (2016) extended their former study (Peters et al., 2015) with the inclusion of Altmetric.com. Their results confirmed that PlumX is the best provider for covering non published materials such as research data. The most recent comparative approach was performed by Meschede and Siebenlist (2018), who compared Altmetric.com and PlumX, finding that less than half of the publications analyzed are included in Altmetric.com, while PlumX covers almost the totality (99%). Zahedi and Costas (2018) performed the most exhaustive comparison between data providers, finding substantial differences in the metrics offered by these platforms.

Many of these studies have focused more on coverage of altmetrics than on counting events and, in the case that events were studied, they were analyzed at research area or source level. This study thus attempts to investigate these counting differences comparing the altmetrics of each article in three representative sources. We consider that this method would bring more detailed results.

Objectives

The main purpose of this study is to compare the results of three important altmetrics providers of the same set of publications. The objective is to determine the differences and similarities between services when they extract and gather these metrics, detecting inconsistencies between different results. These outputs would allow the scholarly community to value the quality and reliability of these services as data providers. Several research questions were formulated:

- What are the differences between data providers according to the proportion of altmetric events captured?
- What are the counting inconsistencies between the metrics provided by these altmetrics services? Is it possible to quantify these differences?
- Could these altmetric providers be used for research evaluation equally?

Methods

Three of the most important altmetric providers were selected for this study: Altmetric.com, PlumX and Crossref Event data (CED). Two main criteria were applied to select these platforms. First, altmetric providers should have a global coverage, not limited to specific publishers, disciplines or regions. In this way, altmetric services of publishers (i.e. PLOS, Nature Publishing Group, etc.) were excluded. Second, these services should offer possibilities to extract data from them, using a public API or scraping their web pages. Due to these restrictions, ImpactStory was not analyzed because it does not provide an API key (ImpactStory, 2017). Both conditions are necessary to be fulfilled in order to make a fair comparison among platforms.

Altmetric providers

PlumX: PlumX (plu.mx/plum/g/samples) is a provider of alternative metrics created in 2012 by Andrea Michalek and Michael Buschman from Plum Analytics. This product is addressed to the

institutional market, offering altmetric counts of publications for particular institutions. PlumX is the aggregator that offers more metrics, including citation and usage metrics (i.e. Views and Downloads). It covers more than 52.6 million of artifacts, being the largest altmetric aggregator (Plum Analytics, 2018). However, only a limited set of institutions make their profiles public, therefore, not all the publications tracked by PlumX are accessible. In 2017, Plum Analytics was acquired by Elsevier (www.elsevier.com) and its altmetric information was added to the Scopus database (Elsevier, 2017). This change could influence data sources, because many of their sources are EBSCO services (www.ebsco.com), the former proprietary.

Altmetric.com: It was the first altmetric provider and it was initiated in 2011 by Euan Adie, with the support of Digital Science (www.altmetric.com). Unlike PlumX, Altmetric.com is centered in the publishing world, signing agreements with publisher houses to monitor the altmetric impact of their publications. This information is accessible through a public API. Today, Altmetric.com tracks the social impact of close to 9 million of research papers (Altmetric.com, 2018). However, this platform does not include metrics about citations and usage in its public API. In 2015, Altmetric.com launched Altmetric for Books, an exclusive service for books and book's chapters (King, 2015).

Crossref Event data (CED): CED is the youngest service, created in 2016 and officially released in 2017. Due to this, the platform claims that the service is still in beta (www.crossref.org/services/event-data). Unlike the previous ones, CED is not a commercial site and it provides free access to their data through a public API. Another important difference is that it does not provide metrics, but it only displays information about each altmetric event linked to a DOI identifier. For instance, it shows the information about the mention of an article on Twitter (date, user, tweet, etc.), but it does not show a count of the number of tweets. For that reason, CED's data would have to be processed to be comparable with the other services.

Data extraction

An initial sample of 67,147 research articles was obtained from the public pages of PlumX. This platform was used as the starting source because it does not permit to retrieve publications by searching by DOI. This drawback makes the search of publications from other altmetric sources in the PlumX database impossible. Another important reason is that this platform only shows altmetric information of publications from institutions that make their profiles public. Therefore, we may suppose that those publications could also be indexed in the other services because they have no problem to display their records. Another reason is that PlumX includes both publications that have and that do not have altmetric events, which also allows us to detect articles that may have altmetric events in other aggregators.

The first step was to identify the organizations that make their profiles public. A search in Google (site:plu.mx) retrieved the web pages of 17 research institutions. The sites that contribute most publications to the sample are: *Mount Sinai Health System* (19,827), *Concytec* (8,141), *University of Pittsburgh* (8,103), *Seoul National University, College of Medicine* (7,871), *King Abdullah University of Science and Technology* (7,281), *University of Helsinki* (5,287), *Xi'an Jiaotong University* (2,781), *International Islamic University of Malaysia* (2,350), *Universidad*

del Rosario (1,648), *Georgia Southern University* (1,083), and *Centro Hospitalar e Universitário de Coimbra* (1,009). The remaining institutions contribute with fewer papers. Only article types published after 2013 and with a DOI were selected, being in total 67,147 research papers. This selection process could introduce a bias because the articles were not randomly gathered and the sample was selected from specific organizations. Later on, we will discuss this possible limitation with previous studies.

This initial sample of research papers was searched again in Altmetric.com using its public API (api.altmetric.com), and in CED using its API (query.eventdata.crossref.org) as well. In this way, three samples were obtained that were comparable amongst the providers. A routine in SQL was written to crawl and extract the altmetric information from the three providers. The samples were obtained in May 23rd, 2017.

Metrics

Because each service has different numbers and types of metrics, it is necessary to define the metrics analyzed in this study. These indicators have to be common to the three platforms and they have to be counted in a similar way. Five metrics were compared:

- Blog posts: number of blog posts that mention a research article. Each service has its own list of blogs where the mentions are searched. Differences between services could be caused more by the number of indexed blogs than by the way in which the mentions are extracted. The only service that provides a number of blogs is Altmetric.com, which monitors 11,000 blogs (Williams, 2017). The other platforms do not make public the list of blogs nor the number of blogs covered.
- News: number of news feeds that mention a research article. In the manner of blog posts, each platform has their own list of news media. Therefore, differences between services could also be due to coverage criteria. Altmetric.com collects a list of 1,300 news outlets, which could be expanded to 80,000 thanks to a partnership with Moreover.com (A LexisNexis Company) (Williams, 2015). The number of news media covered by PlumX is unknown. CED defines this section as “web” and it includes links to other sites different from blogs, thus it is possible that many of these counts do not belong to news mentions.
- Tweets: number of tweets that cite a research article. In this case, the source is the same for the three sites. Possible differences might be found in the way that they obtain their data (i.e. data providers) and how they search the article in the source (i.e. use of identifiers). It is important to mention that Altmetric.com shows only the number of individual accounts (tweeters) that tweet or retweet a publication, while PlumX and CED include all the tweets that mention a publication.
- Wikipedia citations: number of Wikipedia entries that include a citation to a research article. This is a unique source as well, and the differences might be caused by the way in which the citations are extracted.
- Mendeley readers: number of users in Mendeley that include a research article in their libraries. As the two previous metrics, differences in this count would be caused by technical reasons.

Results

Distribution of altmetrics by provider

Table 1 presents the number of documents in the three samples that have altmetric events. Overall, Altmetric.com includes 28,123 articles (41.84%) of the initial sample of 67,218 papers in PlumX. This proportion shows that not even one out of every two articles has statistics in both services, making evident a low overlapping rate. According to CED, the overlap is much higher because 63,637 articles (94.67%) of the PlumX's sample are as well registered in CED. However, only 2,635 (3.92%) of these documents have altmetric events. This absence of data could be due to the service is still in beta and many of the events are not counted yet.

	PlumX	PlumX %	PlumX (altmetrics) %	Altmetric.com	Altmetric.com %	CED	CED %
Blogs	1,787	2.66	3.37	3,334	11.86	37	1.40
News	852	1.27	1.61	4,493	15.98	132	5.01
Tweets	11,803	17.56	22.28	25,699	91.38	896	34.00
Wikipedia	1,168	1.74	2.21	841	2.99	844	32.03
Readers	52,256	77.74	98.66	26,930	95.76	0	0
Total_records	67,218			28,123		2,635	

Table 1. Number of articles with at least one altmetric event by provider

Observing the distribution of altmetric events in the three services, one can detect that there are important differences according to the percentage of articles with an altmetric event. PlumX, the original sample, shows a high proportion of articles bookmarked in Mendeley (77.74%) and mentioned in Twitter (17.56%), and a lower proportion mentioned in blogs (2.66%), news (1.27%) and Wikipedia (1.74%). However, in the case of Altmetric.com, these percentages are much higher. Thus, the proportion of articles saved in Mendeley is 95.76% and mentioned in Twitter is 91.38%. These high percentages are due to Altmetric.com only recording articles that have at least one altmetric event, whereas PlumX indexes articles without distinction. If the articles without altmetrics (14,251) are subtracted, PlumX would then include more articles with readers (98.66%) than Altmetric.com (95.76%). However, this subtraction does not explain why the proportion of tweeted papers in PlumX (22.28%) is much lower than in Altmetric.com (91.38%). This significant small proportion of tweets was already noticed by Jobmann et al. (2014) and this could be due to the fact that PlumX did not use Gnip (support.gnip.com/sources/twitter/), the official Twitter data provider, until 2016.

Regarding CED, the number of articles with an altmetric event is very low, with only 2,635 papers (3.92%) from the initial PlumX's sample. This demonstrates that CED is still in beta and the coverage of altmetric events is for now incomplete. Taking only into account papers with at least one altmetric event, 34% articles have a Twitter mention and 32.03% a citation in Wikipedia. This last is the highest percentage of all the three providers, and it suggests that CED has a special coverage of Wikipedia articles.

Counting differences between providers

The main objective of this work is to analyse the counting differences between altmetric providers. To this end, a new distribution was created from the differences between two altmetric aggregators. For example, distribution Altmetric.com/PlumX shows the counting differences of one metric (blog posts, tweets, citations, etc.) between Altmetric.com and PlumX by each article. The mean of an altmetric distribution (*a*) is calculated by adding the counting differences of each document (*d*) in two providers (*p*), and then dividing them by the number of documents (*D*) with an altmetric event in both providers (*p1* ∩ *p2*). This is an easier and clearer way to present and analyze the differences between providers as well as it allows to detect systematic or random errors. In this way, the dispersion (Standard Deviation, Std. Dv.) informs us about the amount of differences between two services. The more pointed and slender a distribution, the less different the two altmetric providers are to each other.

$$Mean_a = \frac{\sum d_{p1} - d_{p2}}{D_{p1 \cap p2}}$$

Blogs

Figure 1 and Table 2 present the distribution of the counting differences between Altmetric.com, PlumX and CED regarding blog posts. Generally, Altmetric.com is the service that counts more blog posts. In average, 1.52 more posts than PlumX and 2.21 more than CED. These differences suggest that Altmetric.com has the largest list of blogs. In fact, it gathers more than 11,000 blogs. By contrast, the number of blogs that PlumX and CED track is unknown. However, the fact that 13.92% of the articles contain more blog posts in PlumX than in Altmetric.com and that only 10.66% of the articles show the same count in both services would mean that the PlumX's list could be quite different to the Altmetric.com one. It is then possible that PlumX covers an important number of blogs that Altmetric.com does not manage. CED shows more significant differences with respect to the other providers. In almost all the cases, CED contains less blog mentions than Altmetric.com (2.21) and PlumX (1.54). These high differences suggest that the CED's blog list is short and overlapped with the other providers.

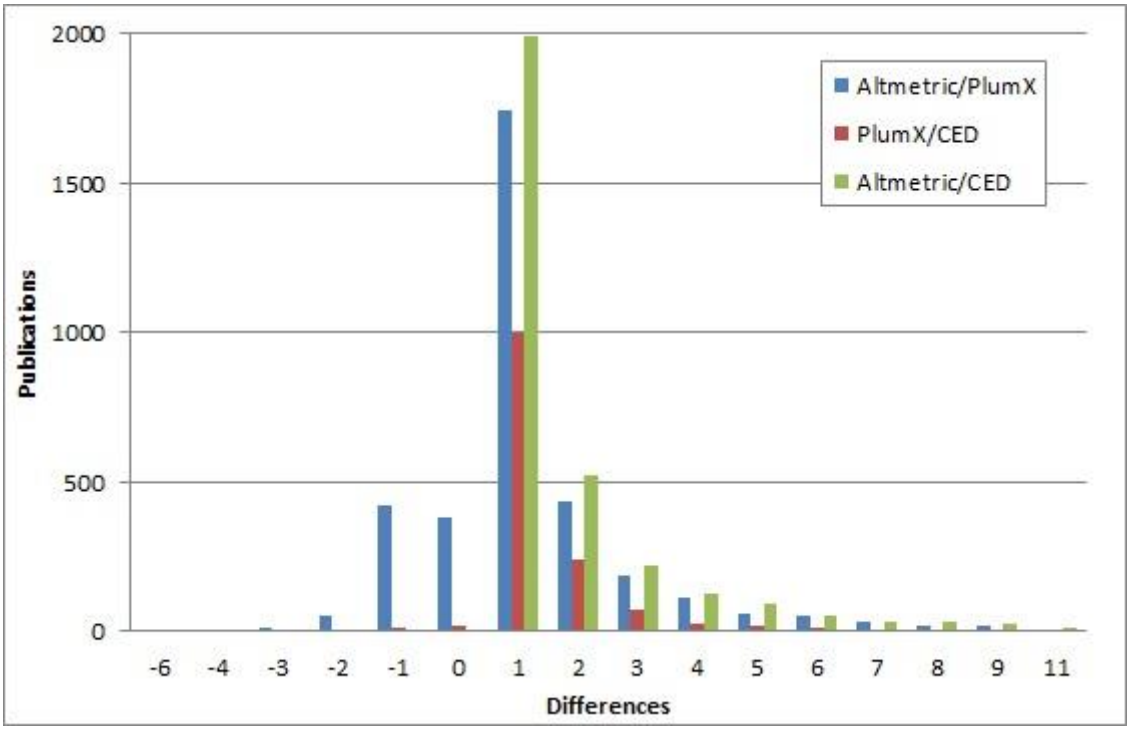

Figure 1. Histograms of counting differences between providers according to blog posts

	Articles	Mean	Std. Dv.	Counting differences			
				<0	0	1	2
Altmetric.com/PlumX	3,585	1.52 (±.09)	2.70	13.92%	10.66%	48.73%	12.13%
PlumX/CED	1,417	1.54 (±.09)	1.79	1.20%	1.13%	70.64%	16.80%
Altmetric.com/CED	3,185	2.21 (±.11)	3.09	0.25%	0.22%	62.48%	16.33%

Table 2. Some parameters of the three comparative distributions according to blog posts

News

Figure 2 and Table 3 show the distribution of the counting differences between providers at the level of news mentions. The number of Articles in Table 2 and following tables do not correspond to the ones reported on Table 1. This is because the Table 1 only includes articles indexed in a platform, whereas tables 2-6 include articles in one or the other provider that have an altmetric event. This fact introduces slight differences in the total counts. Altmetric.com is the service that gathers more comments on news media sites. In average, it collects 6.79 more than PlumX and 6.82 more than CED. These differences cause a high standard deviation of the Altmetric.com's distributions. This fact makes evident that media coverage in Altmetric.com is much wider than in the other services. Some possible causes could be that Altmetric.com can track the mention of articles in the media without a specific identifier (Liu, 2013). Another possible reason could be its partnership with Moreover.com, a news data provider, which allows Altmetric.com to collect more news media (Williams, 2015). According to CED, it is possible that many of these counts do not exactly belong to news mentions. Therefore, the percentage of real news mentions could be even lower. Differences between PlumX and CED are smaller (1.37), which could mean that both services have similar news sources.

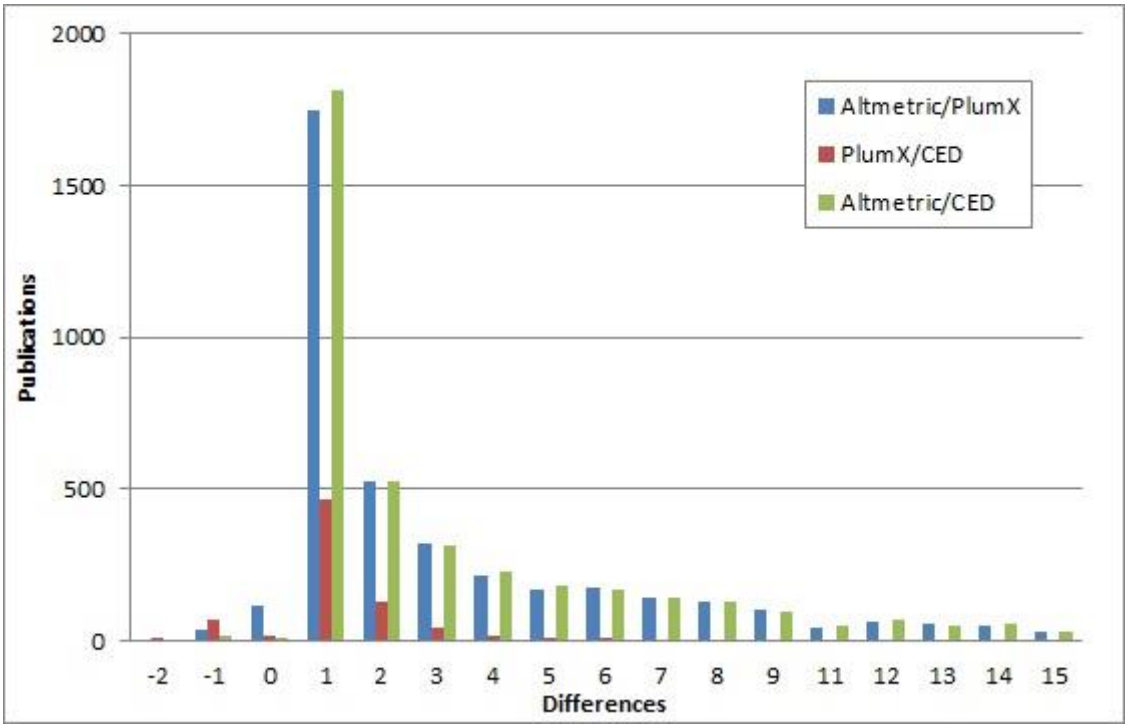

Figure 2. Histograms of counting differences between providers according to news

	Articles	Mean	Std. Dv.	Counting differences			
				<0	0	1	2
Altmetric.com/PlumX	4,224	6.79 (±.45)	14.95	0.92%	2.63%	40.30%	12.08%
PlumX/CED	791	1.37 (±.17)	2.38	0.50%	0.23%	41.90%	12.23%
Altmetric.com/CED	4,314	6.82 (±.43)	14.46	12.00%	2.23%	57.97%	15.70%

Table 3. Some parameters of the three comparative distributions according to news

Tweets

Figure 3 and Table 4 present the distribution of the counting differences between Altmetric.com, PlumX and CED regarding Twitter mentions. On average, Altmetric.com again counts 3.72 more tweets than PlumX and 9.21 more than CED. It is interesting to notice that there is a slight symmetry between Altmetric.com and PlumX. 13.07% of the articles have more tweets in PlumX than in Altmetric.com, 11.34% show the same number of mentions in both platforms and 35.39% have one more tweet in Altmetric.com than in PlumX. This symmetry could be due to Altmetric.com only counting tweeters, different tweets from the same account are hence not counted. By contrast, PlumX does count those tweets. This could explain why there are 13.07% of articles that have more tweets in PlumX than in Altmetric.com. However, the remaining 75.55% of articles with more tweets in Altmetric.com than in PlumX suggest that there is a systematic error. This error could be caused by the employment of different data providers or by the use of distinct criteria in order to count tweets. For example, a possible cause could be that PlumX did not use data from Gnip, the official Twitter data provider, until 2016. This could be the reason why some tweets were not properly matched. A manual inspection of some articles showed that the range of URLs that

PlumX uses to identify mentions is much smaller than those used by Altmetric.com. This absence of completeness would produce the systematic underperforming of PlumX according to Altmetric.com. With regard to CED, it is interesting to notice the great difference with PlumX (14.32) and Altmetric.com (9.21). In this case, there are more systematic variations and these could be due that CED only compiles tweets that contain the DOI prefix or the main landing page of the article, dismissing other links to repositories, publishing platforms, etc (CED, 2017).

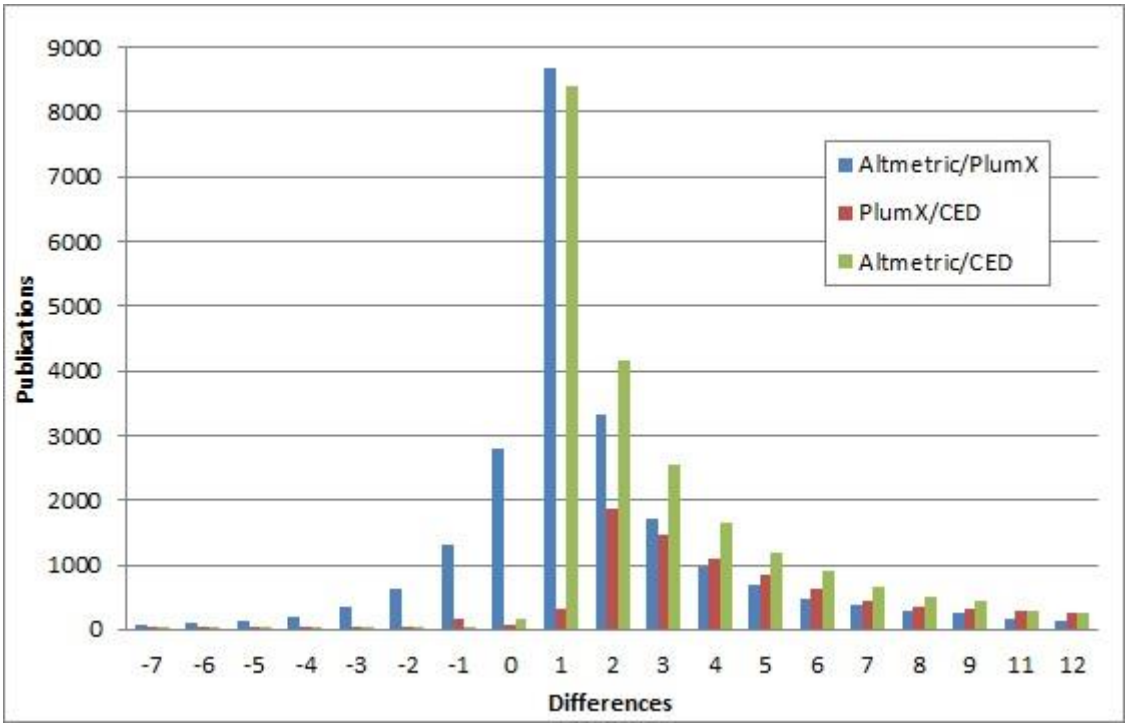

Figure 3. Histograms of counting differences between providers according to tweets

	Articles	Mean	Std. Dv.	Counting differences			
				<0	0	1	2
Altmetric.com/PlumX	21,742	3.72 (±.34)	25.45	13.07%	11.34%	35.39%	13.51%
PlumX/CED	9,951	14.32 (±1.37)	69.69	0.36%	0.72%	34.46%	17.09%
Altmetric.com/CED	24,227	9.21 (±.63)	49.68	3.38%	0.57%	3.17%	18.50%

Table 4. Some parameters of the three comparative distributions according to tweets

Wikipedia

Figure 4 and Table 5 exhibit the distribution of the differences between the three data providers according to Wikipedia citations. In this case, differences are smaller than in the other metrics. This fact is demonstrated by the low averages and the high percentage of articles with the same count value. CED is the platform that gathers the most Wikipedia citations on average, with 0.6 more citations than PlumX and .24 more than Altmetric.com. In addition, PlumX obtains .73 more mentions than Altmetric.com. This metric displays the highest percentages of articles with the same count value, being 66.28% in Altmetric.com/PlumX, 36.66% in PlumX/CED and 36.66% in Altmetric.com/CED. The fact that Wikipedia citations shows such small differences across providers could be caused by the low

number of references in Wikipedia, by the easy way with which the mentions are computed and by the fact that only one source is tracked. In spite of these facts, differences between CED and the other providers are most probably caused by Altmetric and PlumX only tracking mentions on the English, Finnish and Swedish-language Wikipedias (Altmetric Support, 2017), while CED explores all languages on Wikipedia. Another possible reason is that Altmetric.com does not extract mentions from incomplete references or without an identifier. However, PlumX counts mentions only matching the title of the article (Plum Analytics, 2016). This difference would explain the slightly higher number of mentions computed in PlumX according to Altmetric.com.

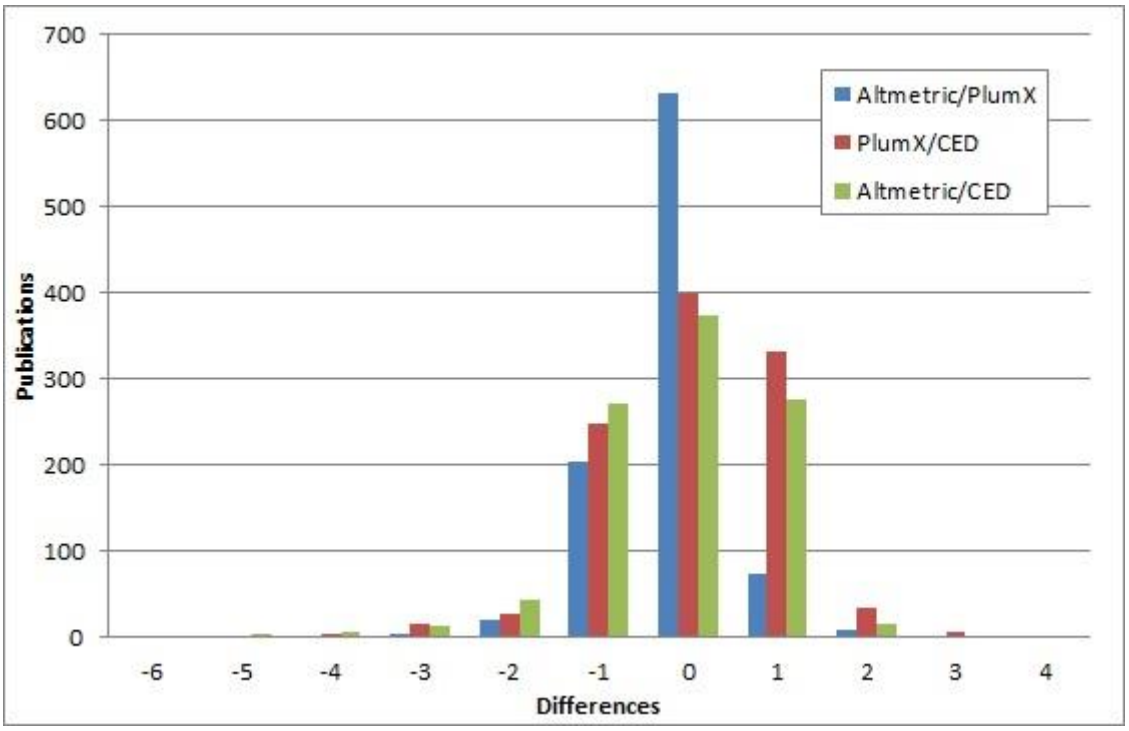

Figure 4. Histograms of counting differences between providers according to Wikipedia citations

	Articles	Mean	Std. Dv.	Counting differences			
				<0	0	1	2
Altmetric.com/PlumX	955	-0.73 (±.21)	1.95	24.92%	66.28%	7.75%	0.94%
PlumX/CED	1,091	-0.6 (±.33)	3.31	28.69%	36.66%	30.52%	3.21%
Altmetric.com/CED	1,023	-0.24 (±.25)	4.29	34.31%	36.66%	26.98%	1.66%

Table 5. Some parameters of the three comparative distributions according to Wikipedia citations

Readers

In the case of Mendeley readers, only Altmetric.com and PlumX include this altmetric indicator. CED does not offer information about this reference manager. Figure 5 depicts the distribution of the differences between Altmetric.com and PlumX compiling Mendeley readers, and Table 6 shows some parameters about this distribution. The histogram is nearly symmetric, with an important proportion of similar counts (28.28%). However, there is a

slightly higher proportion of negative values, which means that PlumX regularly counts more readers than Altmetric.com. In fact, the mean of the differences (-2.5) confirms this counting error. This result is significant because we are talking about only one source and the metric is already counted by the source, that is, providers do not have to aggregate readers. This error is explained by the fact that Mendeley allows to include references that were not previously registered in its Public Catalog. The risk of this practice is that it creates duplicate records. In this case, PlumX aggregates the readers of possible duplicate records from references with similar titles, years, authors, etc. Contrarily, Altmetric.com only counts the readers from a unique identifier, dismissing readers from duplicate items. This could be the reason why on average, PlumX shows more readers than Altmetric.com.

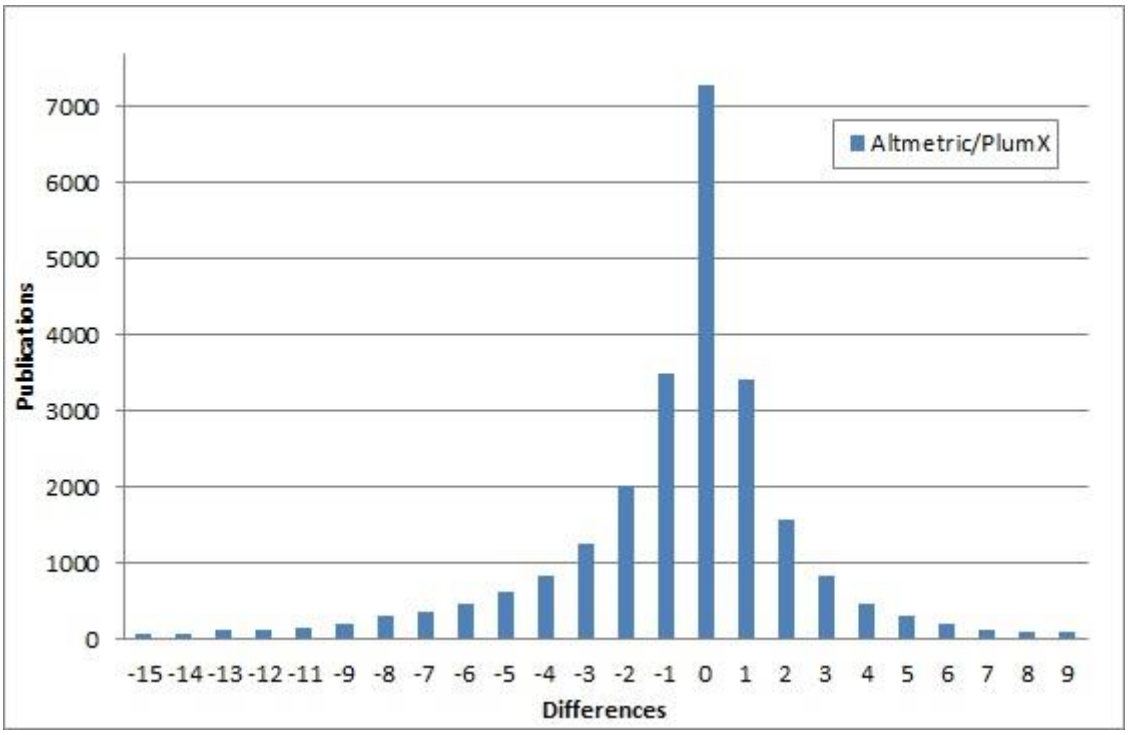

Figure 5. Histograms of counting differences between Altmetric/PlumX according to Mendeley readers

	Articles	Mean	Std. Dv.	Counting differences			
				<0	0	1	2
Altmetric.com/PlumX	18,436	-2.5 (±.48)	33.42	42.35%	28.28%	13.32%	6.07%

Table 6. Some parameters of the comparative distribution of Altmetric.com/PlumX according to Mendeley readers

Discussion

The results show that there are important differences between providers when they collect and count altmetrics. In general, Altmetric.com captures more altmetric impact than the other services. Except for the number of Mendeley readers and Wikipedia citations, Altmetric.com exhibits the highest percentage of papers with at least one altmetric event for each metric. Perhaps, the most significant difference is regarding the coverage of Twitter. Almost all the articles in altmetric.com have been mentioned on Twitter (91.38%), a slightly higher

proportion than that found by Robinson-García et al. (2014). Meanwhile, only 34% of articles in CED and 22.28% in PlumX were tweeted. This last proportion is similar to the one found by Meschede and Siebenlist (2018) (21%) and it was already noticed by Jobmann et al. (2014), mainly with regards to PlumX. This variation could be due to the fact that this aggregator did not use Gnip, the official Twitter data provider, until 2016. Another reason could be that the number of identifiers used by PlumX and CED is lower than the ones used by Altmetric.com, causing a reduction of the number of captured tweets. Altmetric.com also has a better coverage of blogs and news, with more than twice as many mentions. However, there are some metrics in which PlumX and CED perform better than Altmetric.com. For instance, PlumX (98.66%) covers more articles with Mendeley readers than Altmetric.com (95.76%); and CED captures more Wikipedia citations (32%) than Altmetric.com (2.99%) and PlumX (2.21%). These coverage differences allow us to imply that, despite Altmetric being the comparative best provider, it would be recommended to select particular providers to analyze specific metrics. Thus, while Altmetric.com would be the best source for news, blog posts and tweets, PlumX would be more suitable to study Mendeley readers and CED to study Wikipedia mentions.

According to counting differences, the results also show that there are important variations between providers in the five metrics analyzed. Altmetric.com is the service that has a better coverage of blogs, news and tweets, a fact also confirmed by Baessa et al. (2015). In these metrics, Altmetric.com shows more news mentions than PlumX and CED, being more visible in news feeds, with 6.79 more news than PlumX and 6.82 more than CED. This result makes it clear that Altmetric.com is more concerned with having a complete list of blogs and news, which lets it capture the widest media impact possible. Altmetric.com also gathers more tweets than their competitors (3.72 more than PlumX and 9.21 more than CED), caused perhaps by the use of more unique identifiers such as DOIs, repository handles, landing pages, etc.

By contrast, PlumX achieves a better coverage of Mendeley readers (2.5 more) and Wikipedia citations (.73 more) in comparison to Altmetric.com, although these differences are smaller and probably due to technical criteria. For example, Altmetric.com obtains fewer Mendeley readers because it fails to aggregate duplicate records. In comparison with CED, PlumX has better figures in all metrics, except for Wikipedia citations (.6 fewer than CED). This difference is especially significant when counting tweets, where PlumX captures 14.32 more tweets than CED.

These results have also indicated that CED is still a beta service. The number of collected events is rather low in comparison with the other providers, whereas this difference is more noticeable in Altmetric than it is in PlumX. This fact is evident in all the metrics, with the only exception of Wikipedia citations. CED covers all Wikipedia languages, collecting slightly more citations than Altmetric.com (.24) and PlumX (.6).

These results have strong implications for research evaluation because they demonstrate that the counts are not uniform across providers. This means that the altmetric impact of scientific publications is different according to the provider used and it casts doubts on the reliability of these tools for measuring altmetric impact. These differences are mainly due to systematic errors caused by coverage problems (blog posts, news, Wikipedia citations) or technical

limitations (tweets, Mendeley readers). Such results recommend using specific providers if we want to track particular metrics. In this sense, Altmetric.com is the most suitable tool for blogs, news and tweets, PlumX for Mendeley readers and CED for Wikipedia citations. On the contrary, if we want to analyze several metrics, these results, then, suggest to use several aggregators. In this way, altmetric studies could be more consistent selecting the most appropriate sources depending on the metric to be analyzed.

This study could be limited by the sample gathering process. Data were not randomly extracted from PlumX, which could influence the results about the coverage of Altmetric.com and CED in comparison to PlumX. However, this fact would be less relevant when the providers are compared according to counting differences. This is because these counting differences are independent of the way in which the sample was taken. This statement is reinforced when the obtained figures are comparable to previous studies.

Conclusions

There are differences in the amount of articles with altmetric impact in each provider. Altmetric.com is the service with the highest percentage of publications with some altmetric event, surpassing the other providers in almost all the indicators. The only exception is Wikipedia citations where CED obtains very high ratios and in Mendeley readers where PlumX slightly exceeds it. Otherwise, PlumX yields a significantly low proportion of tweeted papers and CED only stands out with extracting Wikipedia citations.

There are also important differences between providers when it comes to altmetric counts. Overall, Altmetric.com is the platform that collects more altmetric events, obtaining the best scores on blog posts, news and tweets. PlumX is only better than Altmetric.com when it compiles Mendeley readers, but it is limited when it comes to extracting blogs, news and tweets. CED is still a beta product with a short range of altmetrics and a deficient collection of news and tweets. However, it is the best service for extracting Wikipedia citations.

Finally, these results suggest that Altmetric.com, PlumX and CED cannot be used separately for research evaluation because there are important differences according to coverage and counting procedures. Therefore, it is recommended to select particular providers for the analysis of specific metrics (i.e., CED for Wikipedia citations, PlumX for Mendeley readers) or to combine several sources if one wants to perform a global analysis about the interaction of several altmetrics.

References

Adie, E. & Roe, W. (2013). Altmetric: enriching scholarly content with article-level discussion and metrics. *Learned Publishing*, *26*(1): 11-17. DOI: 10.1087/20130103

Altmetric.com (2018). How it works. https://www.altmetric.com/about-our-data/how-it-works/

Altmetric Support (2017). How does Altmetric track mentions on Wikipedia? https://help.altmetric.com/support/solutions/articles/6000060980-how-does-altmetric-track-mentions-on-wikipedia- [Visited: May 23, 2018]

Baessa, M., Lery, T., Grenz, D. & Vijayakumar, J. K. (2015). Connecting the pieces: Using ORCIDs to improve research impact and repositories. *F1000Research*, *4:* 195. DOI: 10.12688/f1000research.6502.1

Bornmann, L. (2014). Validity of altmetrics data for measuring societal impact: A study using data from Altmetric and F1000Prime. *Journal of Informetrics*, *8*(4): 935-950. DOI: 10.1016/j.joi.2014.09.007

Chamberlain, S. (2013). Consuming article-level metrics: Observations and lessons from comparing aggregator provider. *Information Standards Quarterly*, 25(2): 4-13.

Champieux, R. (2015). PlumX. *Journal of the Medical Library Association*, *103*(1): 63. DOI: 10.3163/1536-5050.103.1.019

CED (2017). Twitter – Event Data User Guide https://www.eventdata.crossref.org/guide/sources/twitter/ [Visited: May 23, 2018]

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*(10): 2003-2019. DOI: 10.1002/asi.23309

King, G. (2015). Altmetric and Springer Launch Bookmetrix: A New Platform for Book Impact. Digital Science News Blog. https://www.digital-science.com/blog/news/altmetric-and-springer-launch-bookmetrix-a-new-platform-for-book-impact/ [Visited: May 23, 2018]

Elsevier. (2017). Elsevier Acquires Leading 'Altmetrics' Provider Plum Analytics. https://www.elsevier.com/about/press-releases/corporate/elsevier-acquires-leading-altmetrics-provider-plum-analytics [Visited: May 23, 2018]

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(2), 1117-1166. DOI: 10.1007/s11192-016-2077-0

Fraumann, G., Zahedi, Z. & Costas, R. (2015). What do we know about Altmetric.com sources? A study of the top 200 blogs and news sites mentioning scholarly output. Altmetrics workshop, 9 October 2015, Amsterdam Science Park, Amsterdam.

Gumpenberger, C., Glänzel, W. & Gorraiz, J. (2016). The ecstasy and the agony of the altmetric score. *Scientometrics*, *108*(2): 977-982. DOI: 10.1007/s11192-016-1991-5

Holmberg, K. J. (2015). *Altmetrics for information professionals: Past, present and future*. Chandos Publishing. ISBN: 9780081002773

ImpacStory (2017). New widget and API. ImpactStory blog. http://blog.impactstory.org/new-widget-and-api/ [Visited: May 23, 2018]

Jobmann, A., Hoffmann, C.P., Künne, S., Peters, I., Schmitz, J. & Wollnik-Korn, G. (2014). Altmetrics for large, multidisciplinary research groups: Comparison of current tools.

Bibliometrie – Praxis und Forschung, 3, http://www.bibliometrie-pf.de/article/viewFile/205/258

Kraker, P., Lex, E., Gorraiz, J., Gumpenberger, C. & Peters, I. (2015). Research data explored II: The anatomy and reception of figshare. *arXiv preprint arXiv:1503.01298*.

Lindsay, J. M. (2016). PlumX from Plum Analytics: Not Just Altmetrics. *Journal of Electronic Resources in Medical Libraries*, *13*(1): 8-17. DOI: 10.1080/15424065.2016.1142836

Liu, J. (2013) New Feature: News Tracker. Altmetric: https://www.altmetric.com/blog/new-feature-news-tracker/ [Visited: May 23, 2018]

Meschede, C., & Siebenlist, T. (2018). Cross-metric compatability and inconsistencies of altmetrics. *Scientometrics*, *115*(1), 283-297. DOI: 10.1007/s11192-018-2674-1

Ortega, J. L. (2016). To be or not to be on Twitter, and its relationship with the tweeting and citation of research papers. *Scientometrics*, *109*(2): 1353-1364. DOI: 10.1007/s11192-016-2113-0

Peters, I., Kraker, P., Lex, E., Gumpenberger, C., & Gorraiz, J. (2015). Research data explored: Citations versus altmetrics. In A.A. Salah, Y. Tonta, A.A. Akdag Salah, C. Sugimoto, & U. Al (Eds.), *Proceedings of the 15th International Conference on Scientometrics and Informetrics, Istanbul, Turkey* (pp. 172–183)

Peters, I., Kraker, P., Lex, E., Gumpenberger, C. & Gorraiz, J. (2016). Research data explored: an extended analysis of citations and altmetrics. *Scientometrics*, *107*(2): 723-744. DOI: 10.1007/s11192-016-1887-4

Plum Analytics (2016). Wikipedia & Altmetrics: Calculating Mention Metrics. https://plumanalytics.com/wikipedia-altmetrics-calculating-mention-metrics/

Plum Analytics (2018). Coverage: Expanding the World of Altmetrics. khttps://plumanalytics.com/learn/about-metrics/coverage/

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

Robinson-García, N., Torres-Salinas, D., Zahedi, Z. & Costas, R. (2014). Nuevos datos, nuevas posibilidades: Revelando el interior de Altmetric.com. *El Profesional de la Information*, *23*(4): 359-366. DOI:10-3145/epi.2014.jul.03

Roemer, R. C. and Borchardt, R. (2015). Altmetrics. American Library Association, ISBN: 0838959652

Sugimoto, C. R., 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(9), 2037-2062. DOI: 10.1002/asi.23833

Tattersall, A. (Ed.). (2016). *Altmetrics: a practical guide for librarians, researchers and academics*. Facet Publishing. ISBN: 178330010

Thelwall, M., Haustein, S., Larivière, V. & Sugimoto, C. R. (2013). Do altmetrics work? Twitter and ten other social web services. *PloS one*, *8*(5): e64841. DOI: 10.1371/journal.pone.0064841

Torres-Salinas, D., Gumpenberger, C. & Gorraiz, J. (2017). PlumX as a Potential Tool to assess the Macroscopic Multidimensional impact of Books. Frontiers in Research Metrics and Analytics, 03. DOI: 10.3389/frma.2017.00005.

Trueger, N. S., Thoma, B., Hsu, C. H., Sullivan, D., Peters, L. & Lin, M. (2015). The altmetric score: a new measure for article-level dissemination and impact. *Annals of emergency medicine*, *66*(5): 549-553. DOI:10.1016/j.annemergmed.2015.04.022

Williams, C. (2015) Altmetric to provide expanded mainstream media tracking. Altmetric: https://www.altmetric.com/blog/moreover/ [Visited: May 23, 2018]

Williams, C. (2017) Help Altmetric further expand its coverage: tell us about your favourite blogs. Altmetric: https://www.altmetric.com/blog/expandblogs/ [Visited: May 23, 2018]

Zahedi, Z., Fenner, M. & Costas, R. (2015). Consistency among altmetrics data provider/aggregators: What are the challenges?. The 2015 Altmetrics Workshop, Amsterdam.

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.