

# Metrics and Altmetrics: an exploratory study of a possible correlation between the most cited papers in open and restricted access in 2016-2018

André Pacheco  
Faculty of Arts and Humanities  
University of Coimbra  
Coimbra, Portugal  
andrez.pacheco@gmail.com

Alexandre Sousa  
Faculty of Arts and Humanities  
University of Coimbra  
Coimbra, Portugal  
alemed75@gmail.com

Angela Yanai  
Faculty of Arts and Humanities  
University of Coimbra  
Coimbra, Portugal  
angela\_yanai@ufam.edu.br

Susana Lopes  
Faculty of Arts and Humanities  
University of Coimbra  
Coimbra, Portugal  
susanlopes@gmail.com

Luís Machado  
Faculty of Arts and Humanities  
University of Coimbra  
Coimbra, Portugal  
luismachado@fe.uc.pt

## ABSTRACT

Citations count in reference databases has been consolidated as the traditional method of assessing the impact of a scientific work. However, the recent developments around the diversity of web communication channels triggered the scientific community to start questioning the legitimacy of these metrics as the sole indicators of impact. In this context, alternative metrics based on web indicators have begun to emerge. This study attempts to determine the existence of a correlation between traditional citations and altmetric mentions, while also considering if the type of access — open or restricted — has an influence in the impact of a publication. The study is based on a mixed methodology. The sample was composed by the most cited hot papers extracted from the Web of Science, and the most mentioned papers in Altmetrics, between 2016 and 2018, according to type of access. Their numbers in altmetrics and in reference databases (Web of Science and Scopus) was collected and the data was analyzed using Spearman's correlation coefficient. It is concluded that there is not a significant correlation between citations and altmetrics mentions, and that the type of access is not relevant for a paper's success.

## CCS CONCEPTS

General and reference → Cross-computing tools and techniques  
→ Metrics

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [Permissions@acm.org](mailto:Permissions@acm.org).

TEEM'18, October 24-26, 2018, Salamanca, Spain  
© 2018 ACM. ISBN 978-1-4503-6518-5...\$15.00

DOI:10.1145/3284179.3284224

## KEYWORDS

Altmetrics; Open Access; bibliometrics; spearman's correlation coefficient

### ACM Reference format:

André Pacheco, Alexandre Medeiros, Susana Lopes, Ângela Yanai, Ângela and Luís Machado. 2018. Metrics and Altmetrics: an exploratory study of a possible correlation between the most cited papers in open and restricted access in 2016-2018. In *Proceedings of the 6th International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM 2018) (Salamanca, Spain, October 24-26, 2018)*, F. J. García-Peñalvo Ed. ACM, New York, NY, USA, 7 pages.

## 1 INTRODUCTION

The assessment of the impact of scientific works has been a recurring topic in the discussions and research in the academic communities. Its importance overarches every disciplinary boundary as, at some point in their careers, every professional engaged in scientific production will have to deal with the perception of the relevance of its research.

Traditionally, the evaluation of the impact of scientific research is based on the application of bibliometric indicators that, generally, rely on the count of citations in reference bibliographic databases, and in the calculation of the Journal Impact Factor<sup>1</sup>. The creation of the Science Citation Index, in the 1960s, preceded several studies on the universe of scientific production (at an individual or institutional level, and according to thematic fields), which, ultimately, promoted the assessment of these metrics [1]. Citation has always been the most common form of being acknowledged, therefore its use as the measurement of the impact of a scientific work has been accepted with a seeming ease by the

<sup>1</sup> For a detailed definition of a journal's impact factor, see <http://ipsience-help.thomsonreuters.com/inCites2Live/indicatorsGroup/aboutHandbook/usingCitationIndicatorsWisely/jif.html>.

scientific communities. This indicator is also commonly used as a way of measuring the contribution of the research of an individual or group for the progress of the scientific knowledge [2].

Over the last years, due to the constant growth in the volume of information, especially as a result of the advent of the World Wide Web, this form of assessment has been questioned by some authors/researchers and institutions [3, 4]. This contesting movement and its subsequent quest for new forms of evaluation lies at the heart of the so-called alternative metrics. According to its advocates [5, 6], these metrics facilitate a better understanding not only of the visibility of scientific research in the new platforms and channels used for communicating science, including publishing in open access repositories, but also of the social impact of that research [7]. Many of these ambitions were incorporated by the Declaration on Research Assessment (DORA) and by the Leiden's Manifest, which eventually forged the term "altmetrics" to denote this movement [8].

Priem *et al.*, in the manifest in favor of altmetrics, highlight the following advantages of these metrics: i) a greater versatility, speed and up-to-dateness with the possibility of a daily or weekly automatic collection of data; ii) to allow to analyze the impact of the paper in itself and not of the means used for its dissemination, as well as the impact of those papers outside the academic context; iii) to enable to measure the social scope of influential works that were not peer-reviewed; and iv) the possibility of contributing to the creation of optimized filtering and recommendation systems that act in real time [8].

The potential use of these new types of metrics is enhanced by the growing use of web platforms that support academic activities and research, as well as by the greater number of researchers that participate and interact through the internet in academic networks. Posts, shares, likes, tweets, downloads, bookmarks and comments in different online social media are today an important source of information to track and measure the impact of scientific research [9, 4].

The importance of this movement can be understood by the position of the European University Association whom, in its 2017 report, recommends institutions to encourage their researchers "to publish in OA [Open Access] platforms, including rewards and compliance measures with institutional, national and/or European policies" [10]. These rewards should be made visible by acknowledging indicators and evaluation methods that do not rely exclusively in the impact factor of a journal or in the total sum of citations.

Amongst the discussions regarding altmetrics, it is of particular importance the possible correlation between their results and traditional metrics, with the goal of understanding if they assess similar concepts. Eysenbach [11] and Shuai *et al.* [12] argue there is a strong correlation between the results of both, whereas Thelwall *et al.* [13], as well as Haustein, Costas and Larivière [14], did not find such evidences. These authors [14]

believe that altmetrics should act not as a replacement for traditional metrics in science, but in a complementary way, since they provide a different type of information.

In this framework, we believe that it is relevant to further study these topics by collecting new data that support considerations on the following research questions:

- a. Whether the high attention in certain papers, inferred through a high number of citations in a short period of time, translates into a similar impact in altmetrics.
- b. Whether the papers with a high performance in the new metrics exhibit a similar success in formal contexts, perceived through a proportional number of citations.
- c. Whether the type of access (open or restricted) has any significant influence in the accumulation of citations and/or mentions by papers.

In order to address these questions, an exploratory-comparative study with a mixed methodology was followed. The main goal of this study is to contribute to the debate on the relationship between traditional metrics and altmetrics, as well as on the potential influence of the type of access in defining a paper's successful reception by the scientific community, measured in the number of citations.

Specifically, we formulate the following objectives:

- i. To identify the ten most cited papers in open access (OA) and the ten most cited in restricted access (RA), classified as hot papers in the Web of Science (WoS) and published between 2016 and 2018.
- ii. To identify the ten most cited papers in OA and in RA in the Altmetrics top ranking, published in the same period.
- iii. To annotate the number of citations, according to the data made available by WoS and Scopus, and the mentions in altmetrics of the papers identified in i. and ii.
- iv. To confront and correlate the totals of citations and mentions obtained for each year by the papers identified according to their type of access.
- v. To determine the most dominant type of access in the ranking of the ten most cited and mentioned papers for each year of the study.

## 2 METHODS

The methodology selected to obtain an answer to the previously formulated goals is based on an exploratory study with a monostrand conversion design, which includes aspects of both qualitative and quantitative methodologies [15]. We agree with Punch when he argues that "the methods and data used (quantitative, qualitative or both) should follow from, and fit in

with the question(s) being asked” [16]. Therefore, we opted for a statistical analysis based on a “stratified purposive sampling” [15] collected from the databases WoS, Scopus and Altmetric. An important note needs to be made for the sample regarding 2018. Since the year was still undergoing at the moment of the study, data was available solely up to 29 May 2018. As a result, conclusions for 2018 can only be inferred for this interval.

The data collection was initiated by identifying the highest cited hot papers in WoS. Hot papers were selected as the primary object of analysis due to the fact that they represent papers that received a significant number of citations in a short period after being published<sup>2</sup>. Since recent papers with a higher number of quick citations are those most prone to be influential in alternative communication channels, we believe that the hot papers constitute an adequate object for our study. In this study, we use the term “citations” to refer to the number of formal citations a paper receives in traditional academic databases – WoS and Scopus –, while the term “mentions” is used to express the number of informal citations accumulated by a paper in alternative metrics.

The methodology of this study encompasses two moments. In the first moment, for each individual year of 2016, 2017 and 2018, the titles and the citations numbers were retrieved for the ten most cited hot papers in Open Access (OA) and for the ten most cited hot papers in Restricted Access (RA). Additionally, the rank of the ten highest cited hot papers was also obtained in order to determine the percentage of those in OA. The goal of this procedure was to assess if being available in OA is an important feature for papers to accumulate more citations.

In the next phase, the titles of these 60 papers – 20 per year, ten in OA and ten in RA – were individually browsed in Scopus to first determine if they also occur in that database and, if so, to verify the number of citations. The comparison between the number of citations in both databases has three goals: i) to assess if a difference exists in the visibility of papers according to the database of choice; and ii) to obtain an average value for the number of citations of each paper title, which enhances the consistency of the citations of scientific papers to be analyzed in the following phase. Simultaneously, each of these titles were also queried by the Altmetric plugin<sup>3</sup> in order to retrieve the amount of mentions that they accumulate in alternative metrics. The sources covered by these metrics include public policy documents, mainstream media, online reference managers, post-publication peer-review platforms, Wikipedia, the Open Syllabus Project, patents, blogs, citations in WoS and Scopus, research highlights from F1000, social media (Facebook, Twitter and Google+) and other online platforms such as Reddit<sup>4</sup>.

In the next step, Spearman’s correlation coefficient was used to correlate the average of citations with the corresponding

number of mentions. This coefficient was chosen as it is the most appropriate for our datasets, *i.e.*, nonparametric data with less than 30 occurrences. For each year, the correlation coefficients were computed for the most cited hot papers in OA and in RA. This operation allows to study the potential existence of a correlation between a paper’s citations and the influence echoes it creates in informal communication channels.

In the second moment, a study with a similar framework was carried out in the Altmetric Explorer, in order to complement the data collected for the formal databases. In this step, several searches were made on the Altmetrics database, based on the previous criteria. Therefore, once again, for each year of 2016, 2017 and 2018, we attempted to retrieve the top ten most mentioned papers in OA, the top ten most mentioned papers in RA, and the global most mentioned ten papers. The latest query aims to illustrate the representativity of OA papers in the rank of the most mentioned, in order to determine whether being in OA is important to have a broad dissemination in informal communication channels. However, there was no data available for 2018, as the year is still undergoing. Therefore, the total sample amounts to 40 titles. These were then searched in the WoS and Scopus databases and their number of citations was extracted. The goal of this task was to determine firstly, their availability and, secondly, if a broad dissemination in informal channels can be correlated to a proportionally high number of citations in formal backgrounds. This purpose was operationalized once again through the use of *p*’s correlation coefficient between the average number of citations and the mentions in alternative metrics.

### 3 RESULTS

The first research question in the analysis of the results was to determine if the broad attention drawn by hot papers, represented by a high number of citations in a short span, would be translated into an equally high number of mentions in alternative metrics. The results of the application of Spearman’s correlation to the corresponding data are presented in Table 1, Table 2 and Table 3. Our data shows that the only significant correlation was found in 2018 for papers in RA, with  $p = .001$ . None of the other datasets exhibit a significant correlation, since the value of *p* is higher than .05.

**Table 1: Correlation coefficients between citations and mentions for hot papers in 2016**

| Rank | Open Access       |          | Restricted Access |          |
|------|-------------------|----------|-------------------|----------|
|      | Average citations | Mentions | Average citations | Mentions |
| 1    | 3495              | 70       | 2597              | 1312     |
| 2    | 1698              | 319      | 1116              | 659      |
| 3    | 1628              | 18       | 890               | 209      |
| 4    | 1395              | 936      | 853               | 895      |
| 5    | 1288              | 5        | 804               | 2399     |
| 6    | 1125              | 437      | 629               | 2        |
| 7    | 944               | 1216     | 615               | 262      |
| 8    | 872               | 85       | 586               | 162      |

<sup>2</sup> See: [https://images.webofknowledge.com/images/help/WOS/hp\\_hot\\_papers.html](https://images.webofknowledge.com/images/help/WOS/hp_hot_papers.html).

<sup>3</sup> See: <https://www.altmetric.com>.

<sup>4</sup> For a detailed list of the sources, see <https://www.altmetric.com/about-our-data/our-sources/>.

|              |                           |             |                          |             |
|--------------|---------------------------|-------------|--------------------------|-------------|
| 9            | 813                       | 531         | 586                      | 3794        |
| 10           | 736                       | 608         | 565                      | 58          |
| <b>Total</b> | <b>13994</b>              | <b>4225</b> | <b>9241</b>              | <b>9752</b> |
| <b>Corr.</b> | $r_{sp} = -.47; p = .174$ |             | $r_{sp} = .32; p = .374$ |             |

**Table 2: Correlation coefficients between citations and mentions for hot papers in 2017**

| Rank         | Open Access              |             | Restricted Access        |             |
|--------------|--------------------------|-------------|--------------------------|-------------|
|              | Average citations        | Mentions    | Average citations        | Mentions    |
| 1            | 750                      | 850         | 1657                     | 1660        |
| 2            | 393                      | 13          | 515                      | 58          |
| 3            | 464                      | 1722        | 348                      | 427         |
| 4            | 303                      | 121         | 377                      | 23          |
| 5            | 342                      | 6           | 302                      | 846         |
| 6            | 351                      | 1692        | 195                      | 56          |
| 7            | 295                      | 98          | 245                      | 197         |
| 8            | 261                      | 1           | 220                      | 320         |
| 9            | 320                      | 13          | 220                      | 33          |
| 10           | 296                      | 466         | 98                       | 26          |
| <b>Total</b> | <b>3775</b>              | <b>4982</b> | <b>4177</b>              | <b>3646</b> |
| <b>Corr.</b> | $r_{sp} = .54; p = .111$ |             | $r_{sp} = .43; p = .213$ |             |

**Table 3: Correlation coefficients between citations and mentions for hot papers in 2018**

| Rank         | Open Access              |             | Restricted Access        |             |
|--------------|--------------------------|-------------|--------------------------|-------------|
|              | Average citations        | Mentions    | Average citations        | Mentions    |
| 1            | 74                       | 411         | 39                       | 1590        |
| 2            | 45                       | 775         | 46                       | 901         |
| 3            | 40                       | 102         | 37                       | 7           |
| 4            | 20                       | 1           | 33                       | 3           |
| 5            | 20                       | 149         | 34                       | 4           |
| 6            | 18                       | 3           | 34                       | 872         |
| 7            | 15                       | 29          | 17                       | 1           |
| 8            | 17                       | 0           | 27                       | 2           |
| 9            | 12                       | 85          | 21                       | 0           |
| 10           | 15                       | 2           | 33                       | 0           |
| <b>Total</b> | <b>276</b>               | <b>1557</b> | <b>321</b>               | <b>3380</b> |
| <b>Corr.</b> | $r_{sp} = .59; p = .072$ |             | $r_{sp} = .88; p = .001$ |             |

**Table 4: Correlation coefficients between mentions and citations based on the most mentioned papers in 2016**

| Rank | Open Access |                   | Restricted Access |                   |
|------|-------------|-------------------|-------------------|-------------------|
|      | Mentions    | Average citations | Mentions          | Average citations |
| 1    | 8063        | 154               | 4912              | 276               |
| 2    | 4660        | 2016              | 4319              | 98                |
| 3    | 3753        | 589               | 4297              | 57                |
| 4    | 3020        | 75                | 3735              | 224               |
| 5    | 2958        | 1                 | 3101              | 66                |
| 6    | 2685        | 588               | 3047              | 833               |
| 7    | 2645        | 202               | 2979              | 60                |
| 8    | 2474        | 10                | 2880              | 325               |

|              |                          |             |                           |             |
|--------------|--------------------------|-------------|---------------------------|-------------|
| 9            | 2471                     | 1608        | 2682                      | 244         |
| 10           | 2464                     | 809         | 2516                      | 91          |
| <b>Total</b> | <b>35193</b>             | <b>6052</b> | <b>34468</b>              | <b>2274</b> |
| <b>Corr.</b> | $r_{sp} = .09; p = .803$ |             | $r_{sp} = -.08; p = .829$ |             |

The second question of this study was the complementary endeavor of assessing the potential existence of a relationship between highly mentioned papers in informal metrics and their corresponding impact in scholarly citations. The collected data and the application of Spearman's correlation coefficient are presented in Table 4 and Table 5, as no data is yet available for 2018. These results show a  $p > .05$ , which indicates that a significant correlation between a paper's mentions in the Altmetric website and citations in formal databases has not been found.

**Table 5: Correlation coefficients between mentions and citations based on the most mentioned papers in 2017**

| Rank         | Open Access              |                   | Restricted Access         |                   |
|--------------|--------------------------|-------------------|---------------------------|-------------------|
|              | Mentions                 | Average citations | Mentions                  | Average citations |
| 1            | 4510                     | 100               | 5876                      | 53                |
| 2            | 4281                     | 14                | 5060                      | 8                 |
| 3            | 4016                     | 50                | 4715                      | 58                |
| 4            | 3920                     | 77                | 4410                      | 30                |
| 5            | 3837                     | 14                | 3985                      | Not indexed       |
| 6            | 3526                     | 143               | 3734                      | 47                |
| 7            | 3443                     | 90                | 3641                      | 158               |
| 8            | 3301                     | Not indexed       | 3588                      | 48                |
| 9            | 3301                     | 5                 | 3192                      | 14                |
| 10           | 3281                     | 24                | 3154                      | 178               |
| <b>Total</b> | <b>37416</b>             | <b>517</b>        | <b>41355</b>              | <b>594</b>        |
| <b>Corr.</b> | $r_{sp} = .22; p = .559$ |                   | $r_{sp} = -.27; p = .488$ |                   |

**Table 6: Access type of the most cited and mentioned papers in 2016**

| Rank | WoS       |             | Altmetrics |             |
|------|-----------|-------------|------------|-------------|
|      | Citations | Access type | Citations  | Access type |
| 1    | 3237      | Gold        | 8063       | OA          |
| 2    | 2386      | RA          | 4912       | RA          |
| 3    | 1965      | Gold        | 4660       | OA          |
| 4    | 1628      | Gold        | 4319       | RA          |
| 5    | 1405      | Gold        | 4297       | RA          |
| 6    | 1282      | Gold        | 3753       | OA          |
| 7    | 1065      | RA          | 3735       | RA          |
| 8    | 1041      | Gold        | 3101       | RA          |
| 9    | 888       | RA          | 3047       | RA          |
| 10   | 881       | Green       | 3020       | OA          |

The results regarding the third and final research question of verifying the potential influence of access type in the accumulation of citations and mentions are presented in Table 6,

Table 7 and Table 8. As previously mentioned, no data for 2018 has yet been made available by Altmetrics, so the data for this year only reports to citations observed in WoS. As far as WoS publications are concerned, data shows a dominance of OA papers in 2016 and 2017, with a presence of 7 OA papers out of 10, but only 3 OA papers in 2018 in the top 10. Regarding the performance of most mentioned papers in Altmetrics indicators, we can observe in 2017 that 5 papers are in OA, being the top 3 formed uniquely by RA papers, while in 2016 the number of papers in OA is 4. Amongst these, OA papers occupying the first and third places. This data suggests that there is no significant difference between the access type and number of citations and/or mentions.

**Table 7: Access type of the most cited and mentioned papers in 2017**

| Rank | WoS       |             | Altmetrics |             |
|------|-----------|-------------|------------|-------------|
|      | Citations | Access type | Citations  | Access type |
| 1    | 2473      | RA          | 5876       | RA          |
| 2    | 712       | Gold        | 5060       | RA          |
| 3    | 347       | Gold        | 4715       | RA          |
| 4    | 422       | Green       | 4510       | OA          |
| 5    | 322       | Gold        | 4410       | RA          |
| 6    | 487       | RA          | 4281       | OA          |
| 7    | 332       | Gold        | 4016       | OA          |
| 8    | 349       | Gold        | 3985       | RA          |
| 9    | 337       | RA          | 3920       | OA          |
| 10   | 289       | Green       | 3837       | OA          |

**Table 8: Access type of the most cited papers in 2018**

| Rank | WoS       |             |
|------|-----------|-------------|
|      | Citations | Access type |
| 1    | 70        | Gold        |
| 2    | 46        | RA          |
| 3    | 45        | RA          |
| 4    | 43        | Green       |
| 5    | 38        | Gold        |
| 6    | 37        | RA          |
| 7    | 34        | RA          |
| 8    | 33        | RA          |
| 9    | 32        | RA          |
| 10   | 28        | RA          |

## 4 DISCUSSION

Although it can be sometimes difficult to specify the influences that govern such a complex and volatile phenomenon as papers' citations and mentions in alternative metrics, an effort should be made to interpret and discuss the results.

One of the main concerns of this study was to determine if a correlation exists, between formal citations and informal mentions. Regarding this aspect, based on the data collected up to 29th May 2018, a significant correlation was found only in the case of papers in RA published in 2018. Despite this, we can

observe that the coefficient value for RA papers decreases from  $p = .374$  in 2016 (Table 1) to  $p = .213$  in 2017 (Table 2), and reaching a significant correlation of  $p = .001$  in 2018 (Table 3). Likewise, papers in OA register a similar behavior, ranging from  $p = .174$  in 2016, to  $p = .111$  in 2017 and, finally, to  $p = .072$  in 2018. As a result, with the exception of papers in RA, this data indicates that a highly cited paper will not necessarily be highly mentioned in informal communication channels.

Bearing in mind the notion that it is possible to read a significant correlation when the value of  $p$  is lower than .05, this data may suggest a potential trend of a more significant correlation between citations and mentions over the years. If this behavior remains, it is expected that we can observe in the near future a significant correlation also for the papers in OA. Since there have not been any significant changes in how papers are cited in WoS and Scopus, in the temporal interval of this study, these results may suggest a greater activity in alternative metrics over the years. It is, however, still early to draw any conclusion, as more research on this topic is necessary.

Another focus of our study was to determine if a correlation could also be found between the number of mentions of the ten most mentioned papers in the Altmetric ranking, and the number of times they are cited in WoS and Scopus. The results of our data reveal that no significant correlation was found in any of these coefficients. Therefore, the fact that a paper is highly mentioned in alternative metrics does not necessarily translate into a high number of citations. Despite this, it was possible to observe a similar trend in the value of the coefficients over time, as noted for highly cited papers. In the case of papers in RA, the value of  $p$  registered at  $p = .829$  in 2016 (Table 4) and  $p = .488$  in 2017 (Table 5), whereas for papers in OA we observed that  $p = .803$  in 2016, and  $p = .559$  in 2017.

Regarding the final concern of our study of assessing the degree of the importance of a paper's access type to the number of citations and mentions it receives, our data does not reveal the dominance of one access type over the other, as both papers in OA and RA position well in the most cited rankings (Table 6, Table 7 and Table 8). Additionally, the total sum of citations of papers in RA or OA, per year, does not show any particular favor over OA or RA. For instance, papers extracted from WoS in a RA regime exhibit more citations in 2018 and 2017, but less total mentions in 2017, despite having a higher number of citations. Additionally, amongst the most mentioned papers extracted from the Altmetric rankings, the papers in RA have a combined amount of mentions higher than the OA in 2017, but lower in 2016. This data suggests that the access type is not determinant for the success of a paper — if we understand success as being highly cited or mentioned —, since papers disseminated throughout both channels have been shown to receive a high amount of attention by the community.

## 5 CONCLUSIONS AND FUTURE RESEARCH

Altmetrics still have a long way to go before reaching full consolidation. One of the greatest challenges it faces are the

technical limitations, specifically those concerning the standards for data collection. Altmetric tools are heavily based on unique identifiers for data collection (such as DOI for papers and ORCID for researchers), which significantly hinders the capacity of analyzing a broad range of sources.

A concern also exists about the volatility of the data almetrics are based on. Today we measure the number of views and mentions in social networks such as Facebook and Twitter, and in academic tools such as Mendeley and CiteULike. However, if we observe the landscape of social media fifteen or twenty years ago, none of these services existed. How we will deal with this vulnerability is still a topic that needs further discussion. In addition to this, the discussion of how to measure the value of altmetric collected data is another focal topic. As far as we are aware, there are no methods to define the value of the several types of mentions in social networks. It is hard to say if a mention in Twitter is worth, for example, more than 100 likes in a Facebook post. Therefore, we believe that the analysis of altmetric values can be perceived more in relative terms, by comparison with the values of other inputs, rather than in absolute terms.

Furthermore, the use of Spearman's correlation coefficient to our dataset revealed that a significant correlation between citations and informal mentions was only observed in 2018 in the case of hot papers in RA. Therefore, for the majority of cases, it is possible to conclude that the fact that a paper is highly cited in formal scientific communication channels will not necessarily translate into a high number of mentions in informal backgrounds. Likewise, the papers that harvested the most mentions in alternative metrics are not the most cited papers in WoS and Scopus. Accordingly, we believe that these results show that alternative metrics are not suited to act as a replacement for traditional metrics, as they assess different aspects of a paper's influence, but as a complementary analysis that reveals different perspective on the impact of a scientific publication. For example, in Table 1, papers in ranks 5 [17] and 9 [18] in RA have an abnormally high number of mentions in alternative metrics, whereas their academic citations are on par with other papers. There might be a social explanation for this. Both these studies published in 2016 discuss the effects of the Zika virus, which was a major world public health problem in 2016. It is to be expected that a study on such a sensitive topic at the time could potentially draw more attention on channels like Twitter and Reddit than a less-known theme or field [19]. However, this assumption remains a hypothesis. A dedicated research on the socio-economical drive of almetrics reference counts would be necessary.

Lastly, the analysis of our sample did not reveal any particular benefit in publishing in open access. In fact, papers with a restricted access gathered more combined citations in WoS in 2018 and 2017. In addition, the analysis of the data pertaining the rankings of the ten most cited papers in WoS, per

each year, as well as the most mentioned papers in Altmetrics showed that no form of access dominates these standings. No significant difference was observed amongst the most cited papers in open and restricted access. Therefore, it is possible to conclude that the access type is not a relevant factor in determining the amount of citations and/or mentions received by a paper.

These conclusions are valid for the universe of the data collected. In the future, it would be interesting to add depth to this research, both in terms of time lapse and size. Data for at least the entirety of 2018 should be included once available, and the sample size should be increased for greater consistency of the results. As a future study, one should also take into consideration the journals under analysis. It is possible that some journals are so important in their fields that their subscription becomes nearly mandatory. In the case of papers published in these journals, even if they qualify as restricted access, the majority of researchers should have access to them.

## ACKNOWLEDGMENTS

This study was partially supported by the Portuguese Foundation for Science and Technology, under the PhD research grant (SFRH/BD/131004/2017).

## REFERENCES

- [1] Garfield, E. (1964). Science Citation Index--A new dimension in indexing. *Science*, 144(3619), 649–654. <https://doi.org/10.1126/science.144.3619.649>
- [2] Cronin, B. (1984). *The citation process: The role and significance of citations in scientific communication*. London: Taylor Graham
- [3] Galligan, F., & Dyas-Correia, S. (2013). Altmetrics: Rethinking the Way We Measure. *Serials Review*, 39(1), 56–61. DOI: <https://doi.org/10.1080/00987913.2013.10765486>
- [4] Gouveia, F. C. (2013). Altmétrica: métricas de produção científica para além das citações. *Liinc Em Revista*, 9(1), 214–227. DOI: <https://doi.org/http://dx.doi.org/10.18225/liinc.v9i1.569>
- [5] Thelwall, M. (2008). Bibliometrics to webometrics. *Journal of Information Science*, 34(4), 605–621. DOI: <https://doi.org/10.1177/0165551507087238>
- [6] Priem, J. and Hemminger, B.H. 2010. Scientometrics 2.0: New metrics of scholarly impact on the social Web. *First Monday*. 15, 7–5 (2010). DOI: <https://doi.org/http://dx.doi.org/10.5210/fm.v15i7.2874>.
- [7] Altmetrics: a manifesto: 2010. <http://altmetrics.org/manifesto/>. Accessed: 2018-04-22.
- [8] Altmetrics: A manifesto (v 1.01): 2011. <http://altmetrics.org/manifesto>. Accessed: 2018-04-20.
- [9] Costas, R., Zahedi, Z. and Wouters, P. 2015. The thematic orientation of publications mentioned on social media. *Aslib Journal of Information Management*. 67, 3 (May 2015), 260–288. DOI: <https://doi.org/10.1108/AJIM-12-2014-0173>.
- [10] European University Association 2017. *Towards full Open Access in 2020: aims and recommendations for university leaders and National Rectors' Conferences*.
- [11] 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, 4 (Dec. 2011), e123. DOI: <https://doi.org/10.2196/jmir.2012>.
- [12] Shuai, X., Pepe, A. and Bollen, J. 2012. How the Scientific Community Reacts to Newly Submitted Preprints: Article Downloads, Twitter Mentions, and Citations. *PLoS ONE*. 7, 11 (Nov. 2012), e47523. DOI: <https://doi.org/10.1371/journal.pone.0047523>.
- [13] Thelwall, M., Haustein, S., Larivière, V. and Sugimoto, C.R. 2013. Do Altmetrics Work? Twitter and Ten Other Social Web Services. *PLoS ONE*. 8, 5 (May 2013), e64841. DOI: <https://doi.org/10.1371/journal.pone.0064841>.
- [14] Haustein, S., Costas, R. and Larivière, V. 2015. Characterizing Social Media Metrics of Scholarly Papers: The Effect of Document Properties and Collaboration Patterns. *PLOS ONE*. 10, 3 (Mar. 2015), e0120495. DOI:

- <https://doi.org/10.1371/journal.pone.0120495>.
- [15] Teddlie, C. and Tashakkori, A. 2009. *Foundations of mixed methods research integrating quantitative and qualitative approaches in the social and behavioral sciences*. SAGE
- [16] Punch, K.F. 2013. *Developing effective research proposals*. SAGE Publications
- [17] Mlakar, J., Korva, M., Tul, N., Popović, M., Poljšak-Prijatelj, M., Mraz, J., Kolenc, M., Resman Rus, K., Vesnaver Vipotnik, T., Fabjan Vodusek, V.,
- [19] Boon, C.Y. and Foon, J.W.J. 2014. Altmetrics is an Indication of Quality Research or Just HOT Topics. *Proceedings of the IATUL Conferences*. (Helsinki, Jun. 2014).
- Vizjak, A., Pižem, J., Petrovec, M. and Avšič Županc, T. 2016. Zika virus associated with microcephaly. *New England Journal of Medicine*. 374, 10 (Mar. 2016), 951–958. DOI:<https://doi.org/10.1056/NEJMoa1600651>.
- [18] Rasmussen, S.A., Jamieson, D.J., Honein, M.A. and Petersen, L.R. 2016. Zika Virus and Birth Defects – Reviewing the Evidence for Causality. *New England Journal of Medicine*. 374, 20 (May 2016), 1981–1987. DOI:<https://doi.org/10.1056/NEJMs1604338>.