RESEARCH

To What Extent Does the Open Access Status of Articles Predict Their Social Media Visibility? A Case Study of Life Sciences and Biomedicine

Tahereh Dehdarirad[1] and Fereshteh Didegah[2]

[1] Chalmers University of Technology, Department of Communication and Learning in Science, Hörsalsvägen 2, 41296 Gothenburg, SE

[2] Karolinska Institutet, Stockholm, SE

Corresponding author: Fereshteh Didegah (fereshteh.didegah@ki.se)

This study aimed to determine whether, and to what extent, the OA status and OA type of articles can predict their social media visibility, when controlling for a considerable number of important factors. Those factors, which previous research confirmed their positive association with altmetric counts, were journal impact, individual collaboration, research funding, number of MESH topics, topic, international collaboration, lay summary, being a mega journal, F1000 Score, and gender of first and last authors. The data for this study comprised 83,444 articles and reviews in the research area of Life Sciences and Biomedicine from 2012–2016, retrieved from Medline in November 2018. The results showed that the percentage of OA articles mentioned on altmetric platforms was significantly higher than those of the non-OA articles. Furthermore, Open Access was significantly associated with a higher probability of a paper being mentioned on the studied social media platforms. Compared to non-OA articles, the OA articles had a higher average of tweets, Facebooks posts, news posts, and blog posts. By increase of a unit in the OA status, the average number of tweets, Facebooks posts, news posts, and blog posts increased by 92.7%, 25.7%, 83.9% and 48.4%, respectively. Regarding the OA types (studied as Gold vs non-Gold), our findings showed that the Gold OA articles had a higher average number of Tweets and a higher probability of being mentioned in tweets and blogs.

Introduction

The movement towards providing open access to research outputs was initiated with the Budapest Open Access (OA) Initiative in 2002 (Chan et al. 2002; Piwowar et al. 2018) and it has been widely accepted as a desirable phenomenon and has become a reality in many academic spheres (Kriegeskorte et al. 2012). OA publications are mainly divided into Gold and Green OA. Publications published in a full open access journal are called Gold OA. Green publications are those where a preprint version of the article is made available in, for example, an institutional repository. This could possibly happen after a period of embargo set by the journal that publishes the article (Holmberg et al. 2020). In addition to the broad categories of Gold and Green OA, there are also other types of OA such as Hybrid and Bronze. In the hybrid OA model, publishers publish OA articles in closed-access scholarly journals, after authors have paid article processing charge (APC) (Kanjilal & Das 2015). Bronze publications are those articles made available freely to read on the publisher's website, without an explicit Open license (Piwowar et al. 2018).

This study aims to compare the social media visibility of Open Access (OA) versus non-Open Access (non-OA), and Gold type of OA versus other types of OA articles in Life Sciences and Biomedicine. The visibility of OA vs non-OA articles, in terms of citations, has been widely investigated. Whilst some studies have suggested a causal relationship between article OA status and higher citation counts (Eysenbach 2006; Hajjem et al. 2006; Gargouri et al. 2010), some later studies identified weaknesses in the methodology used in those studies where a citation advantage was found (McCabe and Snyder 2015; Hersh and Plume 2016; Hua et al. 2016). OA articles have also shown to have a social media visibility advantage (McKiernan et al. 2016). A study of over 2,000 articles published in Nature Communications showed that articles published openly received nearly double the number of unique tweeters and Mendeley readers than subscription-based articles (Adie 2014). The results of a similar study on 1,761 Nature Communications articles showed that OA articles received 1.2–1.48 times as much social media attention (Twitter and Facebook) as compared to non-OA articles (Wang et al. 2015). In the area of Information Science, Cintra, Furnival and Milanez (2018) found that OA articles in hybrid journals received more citations, Mendeley readers and Altmetrics attention score, when compared to closed

articles in the same journals. In a study of national publications from Finnish universities (2012–2014), Holmberg et al. (2020) investigated how frequently articles published in national OA journals (Gold journals) received social media attention and citation in comparison to articles in subscription-based journals. The results showed significant disciplinary and platform differences in terms of OA advantage. For example, OA articles in Veterinary Sciences, Social & Economic Geography and Psychology received more citations and social media attention, while the opposite was found for articles within Medicine and Health Sciences.

Few other studies have shed light on factors that affect the social visibility of scientific outputs (Haustein et al. 2015; Andersen & Haustein 2015; Didegah, Bowman & Holmberg 2018). These common factors include the characteristics of articles such as journal impact factor, international collaboration, individual collaboration, funding, abstract readability and abstract title (Didegah, Bowman & Holmberg 2018). The results show that important factors vary across different social media platforms. While institution prestige and country prestige are associated with increased Mendeley readers, blog and news posts, they are insignificant factors for Twitter and Facebook posts (Didegah, Bowman & Holmberg 2018).

Alhoori et al. (2015) studied the impact of OA status of articles on their visibility, whilst controlling for the effects of journal and publication year. The results of the study showed that OA articles received more altmetric counts than non-OA articles on eight online platforms. However, when controlling for journal and publication year, no clear relationship was found between OA and altmetric mentions. In a study of a random sample of 1,000,000 articles published, Fraser et al. (2019) compared the altmetric counts of OA articles categorized in four groups (Gold, Green, Hybrid and Bronze) in comparison to non-OA articles. Their findings showed that the highest altmetric counts were for Bronze and Green OA, and the lowest altmetric counts were for non-OA. However, taking factors such as subject classification and country of authorship into account, the results changed; for example, in Humanities, Gold OA articles were the least tweeted.

Our study follows in the same vein and aims to determine whether, and to what extent, the OA status and OA type (Gold vs. non-Gold) of an article predicts their social media visibility. However, this study extends this line of research by simultaneously controlling for a combination of factors that according to the previous literature were found to be significantly influencing both citation and altmetric counts of research articles. This would help to ensure that the regression model more precisely determines the extent to which OA publishing might influence the online visibility of research articles. Furthermore, many of the factors included in our regression analysis have not been studied in the previous research. These factors include journal impact, mega journals, different types of collaboration, MESH category, multidisciplinarity, research funding, lay summaries, F1000 score and authors' gender. In this paper, we were especially interested in comparing Gold OA to other types of OA (Green, Bronze and Hybrid). The reason for this is Plan S. The Plan S initiative advocates a future where all papers are open access and is based on the view that "free access to all scientific publications from publicly funded research is a moral right of citizens" (European Commission 2019). The goal of the plan is to decrease the gap between the rich and poor, but the results of some studies suggest that this may be difficult due to the APC rate for Gold OA (Van Leeuwen, Robinson-Garcia & Costas 2019).

To address the research objectives, researchers formulated the following questions:

1. To what extent does the OA status (OA, non-OA) of articles predict their social media visibility, when controlling for a number of important factors (journal impact, individual collaboration, research funding, number of MESH topics, international collaboration, lay summary, F1000 Score, mega journal, MESH topic, F1000 Score, and the gender of first and last authors)?
2. To what extent does the OA type (Gold vs. non-Gold) of articles affect their social media visibility, when controlling for a number of important factors (journal impact, individual collaboration, research funding, number of MESH topics, international collaboration, lay summary, F1000 Score, mega journal, MESH topic, F1000 Score, and the gender of first and last authors)?

Methodology
Data collection
The data for this study comprised 83,444 articles and reviews in the research area of Life Sciences & Biomedicine from 2012–2016, retrieved from the Web of Science Medline in November 2018 using this query: *SU = Life Sciences & Biomedicine*. Using articles' PMID, a search was conducted in Altmetric.com (October 2017 version) in order to obtain the following altmetric indicators: tweet counts, Facebook posts, news posts, blog posts, F1000 post counts, F1000 score, policy mentions, and Wikipedia mentions.

Dependent, independent and control variables
The numbers of tweets, Facebook posts, news posts, and blog posts were considered as dependent variables in a regression model (explained below) to measure the extent to which the OA status of an article may predict its social media visibility. OA status of the articles and their OA types were obtained from Unpaywall.org in November 2019. In the same model, the OA status of articles and OA types (gold vs. non-gold OA) were considered as the independent variables and several other variables as control variables or covariates.

The association between OA factors and social media counts may vary by adding or removing controlled factors from the regression model (Fraser et al. 2019). Hence, a number of control variables that may interfere the association were identified and entered into the regression model simultaneously. While previous research has introduced several factors

Dehdarirad and Didegah: To What Extent Does the Open Access Status of Articles Predict
Their Social Media Visibility? A Case Study of Life Sciences and Biomedicine

Art. 5, page 3 of 14

in association with citation counts, not all potential factors are examined in association with altmetric counts. Given the high correlation found between citation and altmetric counts (Thelwall et al. 2013), same citation factors could be also important for social media visibility. Hence, the variables taken into account in this study are the important factors that significantly influenced both citation and social media counts of research articles in previous studies. Journal impact (gauged by SNIP and mega journals in this study) and research collaboration were the most important factors associating with both citation and altmetric counts in difference subject fields (Didegah & Thelwall 2013; Didegah, Bowman & Holmberg 2018). Research funding was found to be an important factor for citation counts in Life Sciences and Medicine (Didegah 2014). Both citation and altmetric counts were also found to vary across subject fields (Didegah & Thelwall 2013; Didegah, Bowman & Holmberg 2018). Because this study is done in the area of Life Sciences and Biomedicine, MESH categories were an appropriate subject classification to consider. Few factors are rather new in the field such as lay summaries, author genders and F1000 score. Lay summaries are a potential factor initially proposed by Didegah, Alperin & Haustein (2018) but its association with altmetric counts is yet a matter of question. Author genders also found to be an important factor on some altmetric platforms (Sotudeh, Dehdarirad, & Freer 2018).

Table 1 lists all these different variables and their descriptions; few variables and how they were measured are further explained below. We also used publication year as an offset variable in the regression model.

To determine the gender of the authors (first and last), we used Gender API (https://gender-api.com/). This service offers a standard first name search with the possibility to handle double names. The response contains gender assignments (male, female, or unknown), plus confidence parameters, samples and accuracy (Santamaría and Mihaljević, 2018). In cases of gender-neutral, unknown, initials or in cases where the accuracy was lower than 80%, the names were checked manually using internet searches and authors' websites. The gender of 35 authors were remained unidentified. In our regression model they were regarded as missing values.

Table 1: Dependent variables, independent variables and covariates for the hurdle model.

Variable type	Name	Measure
Dependent	Twitter	Number of tweets
	Facebook	Number of Facebook posts
	News	Number of news posts
	Blog	Number of blog posts
Independent	OA status	OA (1); non-OA articles (0)
	OA type	Gold OA (1); Other types of OA (green, bronze & hybrid) (2)
Covariate	Individual collaboration	Number of authors collaborating in an article
	International collaboration	International collaboration (1); national collaboration (0)
	SNIP	Source (journal) Normalized Impact per Paper
	Mega journals	Spezi et al.'s (2017) mega journal list: is a mega journal (1): is not (0)
	Funding	Funded article (1); not-funded article (0)
	Multi-topics	Number of MESH headings assigned to each article
	Mesh category	Seven mesh categories assigned to each article as listed below: MESH1: Anatomy (1); Otherwise[1] (0) MESH2: Organisms (1); Otherwise (0) MESH3: Diseases (1); Otherwise (0) MESH4: Chemicals and Drugs; Otherwise (0) MESH5: Analytical, Diagnostic and Therapeutic Techniques and Equipment (1); Otherwise (0) MESH6: Psychiatry and Psychology (1); Otherwise (0) MESH7: Health Care (1); Otherwise (0)
	Authors genders	The gender of first and last authors
	Lay summary	Articles from journals including lay summaries listed in Shailes list[2] (1); other journals (0)
	F1000 score	The score was obtained from Altmetrics.com.

---

[1] By otherwise we mean the other 13 MESH categories.

[2] https://elifesciences.org/articles/25411?utm_source=content_alert&utm_medium=email&utm_content=fulltext&utm_campaign=elife-alerts.

Regarding mega journals as a covariate in the regression model, we used the journal list provided by Spezi et al.'s (2017) study to determine whether a journal was mega journal or not.

Medline assigns articles to 14 broad MESH. In this article, only seven categories including Anatomy, Organisms, Diseases, Chemicals and Drugs, Analytical, Diagnostic and Therapeutic Techniques and Equipment, Psychiatry and Psychology and Health Care as the most relevant medical topics were considered for evaluation. For each of the seven MESH categories, we created a dummy variable as a control variable in the regression models.

F1000 score as an altmetric indicator is included in the model as a control variable. The rationale behind this is that articles scored in F1000 are recommended as highly important works in the fields of life sciences, health and physical sciences and beyond (Faculty Opinions 2020). Thus, this factor may affect the online visibility of articles in the field, regardless of their open access status.

Data analysis and procedures

Hurdle model

Descriptive statistics was used to depict the state-of-the-art of OA articles vs. non-OA articles shared on the seven altmetrics platforms. Two-sample proportion tests were also performed in order to compare the proportion of OA papers shared on different altmetrics platforms in comparison to non-OA articles.

To answer the second research question, given that the dependent variables (altmetric counts) of this study were count data, count regression models were used. As altmetric counts are over-dispersed and include excessive number of zeros, a count model is required to deal with these two issues.

First, a standard negative binomial, a zero-inflated negative binomial and a hurdle negative binomial models were applied. A standard model is frequently used to model overdispersed data. Zero-inflated models are used for overdispersed and excessive zero datasets and assume that there are two types of zeros in the data: zeros which arise from a negative binomial count distribution and zeros which arise from a "perfect-zero" distribution (Hilbe, 2011). Hurdle models measure the likelihood of an observation being positive or zero, and then determine the parameters of the count distribution for positive observations. Thus, a hurdle model comprises two parts: the count model, which is a negative binomial model, and the logit model. The count model predicts changes in non-zero social media counts, whilst the logit model reports the changes in zero social media mentions for a unit change in the open access factors and each of the covariates.

We finally concluded that a negative binomial-logit hurdle model was the best fit for the data as it creates a scenario in which the positive counts follow a Poisson or NB distribution after passing a hurdle to gain positive counts (Didegah, Bowman, & Holmberg 2018).

Therefore, in order to measure the extent to which the OA type (Gold OA vs Green, Bronze, and Hybrid) of an article may affect its social media visibility, researchers ran a a hurdle model to assess the association between OA type of articles and the number of times they were mentioned on each altmetrics platform. The OA type was considered as a binary variable (Gold vs. Non-Gold), because of the small amount of publications that were Green or Hybrid. Moreover, the distribution of Bronze articles on Facebook, news and blogs was too small to be statistically reliable.

Only Twitter, Facebook, news and blogs were considered for the second and third research questions, as there were a significant number of articles visible on these platforms (See **Figure 1**). To obtain the most reliable results, a number of important factors that possibly had an impact on social media visibility of an article were entered into the model as control variables (See **Table 1**).
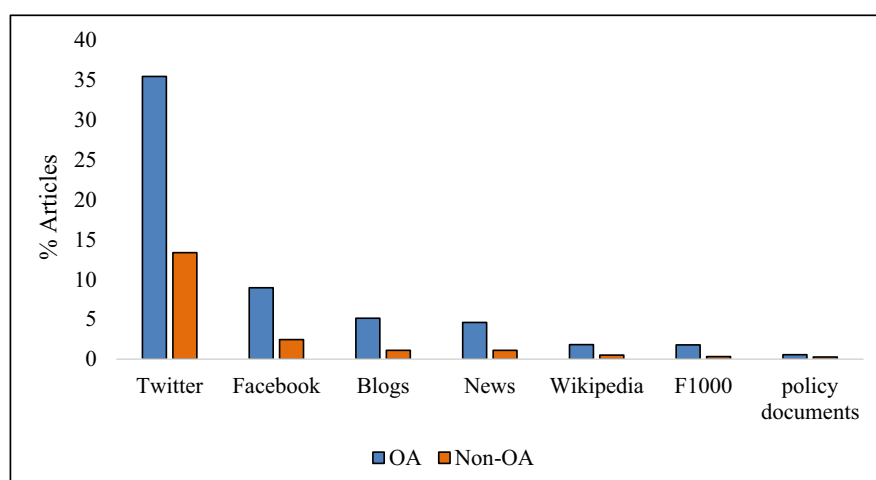

Figure 1: The percentage of OA vs Non-OA articles in Life Sciences and Biomedicine (2012–2016) across different social media platforms.

Dehdarirad and Didegah: To What Extent Does the Open Access Status of Articles Predict
Their Social Media Visibility? A Case Study of Life Sciences and Biomedicine

Art. 5, page 5 of 14

Multicollinearity

Multicollinearity happens when two independent variables/covariates are highly correlated. It is troublesome in regression models because it affects the relationship between the predictors and the output variable (Didegah 2014). A popular test to diagnose multicollinearity is the Variance Inflation Factor (VIF). The VIF is based on the proportion of variance that a predictor variable shares with other predictors in the model. There are several rules of thumb of 4, 10, 20, and over, based on which VIFs over 4, 10, 20, or more are considered to show a high multicollinearity.

The VIF is tested for both the OA status and OA type factors in correlation with the covariates and the results are reported in **Table 2**. As can be seen, the VIFs of both groups of variables are rather low and they do not exceed any standards. That said, this problem will not affect the results of the regression models.

Results

From a total of 83,444 papers, 49,598 (59.44%) were OA and 30,137 (36.12%) were non-OA. 3709 (4.44%) papers did not have any OA status, according to Unpaywall.org. Twitter had the highest percentage of OA articles (35.45%), followed by Facebook (8.95%), blogs (5.13%), news (4.60%), Wikipedia (1.82%), F1000 (1.79%) and policy documents (0.57%), respectively. (See **Figure 1**). The highest percentage of non-OA articles were also shared on Twitter (13.36%), followed by Facebook (2.44%), blogs (1.11%), and news (1.10%).

The results of the two proportion tests also showed that the percentage of OA articles mentioned on altmetrics platforms was significantly higher than that of the non-OA articles [$P < 0.0001$].

Figure 2 shows the percentage of each type of OA. As can be seen from the figure, Gold OA had the highest percentage of articles being followed by Bronze and Green OA types in the second and third place, respectively.

The detailed results of hurdle models for each platform are as follows.

Twitter

Regarding the count model, both the OA vs. non-OA (hereafter OA status) and Gold vs. non-Gold (hereafter OA type) factors were significantly associated with an increase in the estimated number of received tweets. While a unit change in the OA status increased the estimated number of received tweets by 92.7%, the OA type contributed to 28.4% increase in the estimated received tweet counts. Among the types of OA, Gold OA is found to be a more important factor compared to other types of OA. By a unit change in the factor, which means moving from other types of OA together (Green, Bronze, and Hybrid) to Gold OA, the number of tweets to the article on average will approximately increase by 28.4%. The logit model also confirmed that OA status and OA type of articles were significantly associated with the increased probability of a paper being mentioned in Twitter.

Table 2: VIF results for both the OA type and OA status factors.

Variable	VIF	Variable	VIF
OA vs. Non-OA	1.18	Gold vs. Non-Gold	1.93
Indiv. Collaboration	1.27	Indiv. Collaboration	1.28
Int. Collaboration	1.14	Int. Collaboration	1.11
Funding	1.16	Funding	1.15
No. Mesh Categories	1.08	No. Mesh Categories	1.08
Lay summary	1.04	Lay summary	1.3
F1000 score	1.02	F1000 score	1.02
SNIP	1.07	SNIP	1.1
Mega journals	1.12	Mega journals	1.72
First-gender	1.05	First-gender	1.03
Last-gender	1.06	Last-gender	1.03
Mesh category1	1.14	Mesh category1	1.15
Mesh category2	1.22	Mesh category2	1.19
Mesh category3	1.13	Mesh category3	1.13
Mesh category4	1.12	Mesh category4	1.13
Mesh category5	1.29	Mesh category5	1.29
Mesh category6	1.04	Mesh category6	1.04
Mesh category7	1.06	Mesh category7	1.06

Art. 5, page 6 of 14

Dehdarirad and Didegah: To What Extent Does the Open Access Status of Articles Predict
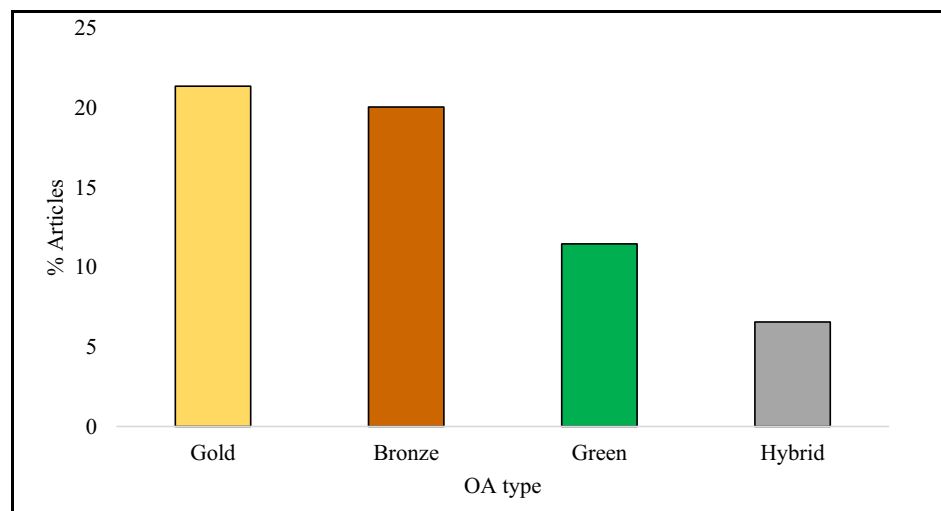Their Social Media Visibility? A Case Study of Life Sciences and Biomedicine

Figure 2: The percentage of articles in Life Sciences and Biomedicine (2012–2016) across different OA types.

Table 3: Hurdle model results for Twitter.

Count model	Coef.	Exp(Coef.)[1]	Sig.[2]	Count model	Coef.	Exp(Coef.)	Sig.
OA vs. Non-OA	0.656	1.927	<2e−16***	**Gold vs. Non-Gold**	0.250	1.284	7.69e−12***
Indiv. Collaboration	0.011	1.011	4.98e−06***	Indiv. Collaboration	0.008	1.008	0.000***
Int. Collaboration	0.278	1.320	<2e−16***	Int. Collaboration	0.312	1.365	<2e−16***
Funding	−0.009	0.991	0.027*	Funding	−0.008	0.992	0.052.
No. Mesh Categories	−0.005	0.995	0.457	No. Mesh Categories	0.001	1.001	0.879
Lay summary	1.190	3.288	<2e−16***	Lay summary	1.000	2.717	<2e−16***
F 1000 score	0.246	1.279	<2e−16***	F 1000 score	0.235	1.265	4.93e−16***
SNIP	0.663	1.940	<2e−16***	SNIP	0.725	2.064	<2e−16***
Mega journals	0.345	1.412	<2e−16***	Mega journals	0.183	1.200	4.10e−05***
First-gender	−0.039	0.962	0.109	First-gender	−0.101	0.904	0.000***
Last-gender	0.042	1.043	0.131	Last-gender	−0.048	0.953	0.136
Mesh category1	−0.428	0.652	<2e−16***	Mesh category1	−0.449	0.638	<2e−16***
Mesh category2	0.047	1.048	0.223	Mesh category2	−0.096	0.909	0.033*
Mesh category3	−0.404	0.667	<2e−16***	Mesh category3	−0.410	0.664	1.96e−14***
Mesh category4	−0.425	0.654	<2e−16 ***	Mesh category4	−0.509	0.601	<2e−16***
Mesh category5	−0.179	0.836	2.66e−10***	Mesh category5	−0.170	0.843	4.58e−07***
Mesh category6	0.394	1.483	8.81e−07***	Mesh category6	0.459	1.583	1.34e−06***
Mesh category7	0.123	1.131	0.044*	Mesh category7	0.125	1.134	0.091.
Logit model	**Coef.**	**Exp(Coef.)**	**Sig.**	**Logit model**	**Coef.**	**Exp(Coef.)**	**Sig.**
OA vs. Non-OA	0.748	2.114	<2e−16***	**Gold vs. Non-Gold**	0.600	1.823	<2e−16***
Indiv. Collaboration	0.033	1.034	<2e−16***	Indiv. Collaboration	0.029	1.029	<2e−16***
Int. Collaboration	0.125	1.133	.52e−13***	Int. Collaboration	0.161	1.175	4.71e−14***
Funding	−0.012	0.988	0.001***	Funding	−0.006	0.994	0.132
No. Mesh Categories	0.030	1.030	.88e−08***	No. Mesh Categories	0.027	1.027	9.93e−05***
Lay summary	1.502	4.491	<2e−16***	Lay summary	1.098	2.998	<2e−16***
F 1000 score	0.329	1.390	<2e−16***	F 1000 score	0.283	1.326	1.28e−13***
SNIP	0.560	1.751	<2e−16***	SNIP	0.737	2.090	<2e−16***

(Contd.)

Dehdarirad and Didegah: To What Extent Does the Open Access Status of Articles Predict
Their Social Media Visibility? A Case Study of Life Sciences and Biomedicine

Art. 5, page 7 of 14

Logit model	Coef.	Exp(Coef.)	Sig.	Logit model	Coef.	Exp(Coef.)	Sig.
Mega journals	0.367	1.444	<2e–16***	Mega journals	−0.051	0.950	0.127
First-gender	0.035	1.036	.035*	First-gender	0.008	1.008	0.711
Last-gender	0.015	1.015	0.432	Last-gender	0.003	1.003	0.894
Mesh category1	−0.146	0.864	.04e–07***	Mesh category1	−0.184	0.832	3.88e–07***
Mesh category2	−0.019	0.982	0.475	Mesh category2	−0.049	0.952	0.136
Mesh category3	0.032	1.033	0.313	Mesh category3	−0.053	0.949	0.194
Mesh category4	−0.112	0.894	.000***	Mesh category4	−0.152	0.859	9.95e–05***
Mesh category5	0.026	1.026	0.192	Mesh category5	−0.003	0.997	0.914
Mesh category6	0.281	1.325	.36e–06***	Mesh category6	0.345	1.412	7.89e–06***
Mesh category7	0.040	1.040	0.356	Mesh category7	0.005	1.005	0.925

[1] Change in the mean parameter of positive citation counts for a unit increase.
[2] Change in the probability of zero citations for a unit increase.
[3] Sig. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

As for the controlled factors in both OA status and OA type models, the count model showed that the journal impact (measured by SNIP), mega journal, individual collaboration, international collaboration, lay summary, F1000 score, and being indexed under 'Psychiatry and Psychology' MESH category significantly contributed to higher number of received tweets, on average. Neither the gender of last author nor the number of mesh categories was found to be an important factor in the OA status and OA types count models (**Table 3**). Furthermore, as demonstrated by the count and logit models for OA status, Funding had a weak negative association with the average number of tweets counts as well as the probability of being mentioned on Twitter.

Facebook

As can be seen from **Table 4**, open access articles had a higher average number of Facebook post counts, when controlling for a number of important factors. A unit change in the OA status (changing from non-OA status to the OA status) approximately increased the average number of Facebook posts by 25.7%. This was the weakest association, when compared with the three other social media platforms. According to the logit models, whilst the open access status significantly associated with a higher probability of being mentioned in Facebook, the open access type did not.

Regarding the controlled factors in count models, international collaboration, the number of MESH topics, F1000 score, and Journal impact (SNIP) were found to be associated with a higher number of received Facebook posts.

By looking at the count models for both OA status and OA types, it can be concluded that funding and 'Anatomy' MESH category were found to be negatively associated with the estimated number of received Facebook posts. Finally, the gender of first and last authors were significant factors for the higher number of received Facebook posts in the OA status model (**Table 4**). However, they were not significant factors for the estimated number of received Facebook posts in the OA type models.

News

As can be seen from both count and logit models for the OA status, for news outlets, the OA status of articles was a significant factor for both a higher probability of a paper being mentioned in news and a higher average number of received news mentions (See **Table 5**). By increase of a unit in the OA factor (changing from non-OA status to OA status), the estimated number of News mentions to articles increased by 83.9%. This association was higher than the association between the received news mentions and international collaboration, journal impact and other significant factors in the model.

However, the OA type was not a significant contributing factor for the estimated number of news mentions or for the probability of being mentioned in news outlets.

For both OA status and OA type models, international collaboration, journal impact (SNIP), number of MESH topics and F1000 score significantly contributed to higher odds of visibility for the articles on News platforms.

Regarding the types of MESH categories, our findings from both logit and count models (OA vs non-OA) showed that articles indexed under specific MESH topics ('Chemicals and Drugs' and 'Analytical, Diagnostic and Therapeutic Techniques') had a news disadvantage in terms of odds of being mentioned in news outlets and the estimated number of received news mentions compared to other MESH topics. Furthermore, the results of logit model for both (OA vs non-OA; Gold vs non-Gold), showed that articles indexed under 'Psychiatry and Psychology' were significantly more likely to be mentioned in news outlets.

Art. 5, page 8 of 14

Dehdarirad and Didegah: To What Extent Does the Open Access Status of Articles Predict
Their Social Media Visibility? A Case Study of Life Sciences and Biomedicine

Table 4: Hurdle model results for Facebook.

Count model	Coef.	Exp(Coef.)	Sig.	Count model	Coef.	Exp(Coef.)	Sig.
OA vs. Non-OA	0.229	1.257	5.52e−05***	**Gold vs. Non-Gold**	0.108	1.114	0.164
Indiv. Collaboration	0.006	1.006	0.106	Indiv. Collaboration	0.004	1.004	0.020*
Int. Collaboration	0.291	1.337	5.95e−10***	Int. Collaboration	0.415	1.515	<2e−16***
Funding	−0.022	0.979	0.002**	Funding	−0.018	0.982	0.012345*
No. Mesh Categories	0.090	1.094	5.63e−11***	No. Mesh Categories	0.087	1.091	1.06e−08***
Lay summary	0.014	1.014	0.863	Lay summary	−0.091	0.913	0.381
F 1000 score	0.155	1.167	7.38e−09***	F 1000 score	0.168	1.183	2.20e−08***
SNIP	0.300	1.350	<2e−16***	SNIP	0.337	1.401	<2e−16***
Mega journals	0.031	1.032	0.638	Mega journals	−0.027	0.973	0.767
First-gender	0.126	1.134	0.008**	First-gender	0.021	1.021	0.699
Last-gender	0.240	1.272	1.45e−05***	Last-gender	−0.020	0.980	0.746
Mesh category1	−0.329	0.720	7.83e−05***	Mesh category1	−0.350	0.705	0.000***
Mesh category2	−0.148	0.862	0.038*	Mesh category2	−0.068	0.934	0.399
Mesh category3	−0.007	0.993	0.936	Mesh category3	−0.043	0.958	0.680
Mesh category4	−0.118	0.889	0.189	Mesh category4	−0.039	0.962	0.701
Mesh category5	−0.130	0.878	0.023 *	Mesh category5	−0.113	0.893	0.080
Mesh category6	0.191	1.210	0.134	Mesh category6	0.275	1.316	0.051.
Mesh category7	0.001	1.001	0.995	Mesh category7	0.161	1.175	0.212
Logit model	**Coef.**	**Exp(Coef.)**	**Sig.**	**Logit model**	**Coef.**	**Exp(Coef.)**	**Sig.**
OA vs. Non-OA	0.733	2.082	<2e−16***	**Gold vs. Non-Gold**	0.027	1.027	0.5
Indiv. Collaboration	0.019	1.019	.77e−15***	Indiv. Collaboration	0.020	1.020	.39e−15***
Int. Collaboration	0.188	1.207	.84e−14***	Int. Collaboration	0.196	1.217	.59e−12***
Funding	0.005	1.005	0.211	Funding	0.004	1.004	0.348
No. Mesh Categories	0.030	1.031	0.000***	No. Mesh Categories	0.026	1.026	0.004**
Lay summary	1.052	2.863	<2e−16***	Lay summary	1.054	2.869	<2e−16***
F 1000 score	0.209	1.232	<2e−16***	F 1000 score	0.187	1.206	.74e−13***
SNIP	0.458	1.581	<2e−16***	SNIP	0.479	1.614	<2e−16***
Mega journals	0.335	1.398	<2e−16***	Mega journals	0.332	1.394	.32e−12***
First-gender	0.045	1.046	0.076.	First-gender	0.025	1.025	0.384
Last-gender	0.044	1.045	0.128	Last-gender	0.031	1.031	0.347
Mesh category1	−0.125	0.883	0.004**	Mesh category1	−0.168	0.846	0.000***
Mesh category2	0.052	1.054	0.173	Mesh category2	0.061	1.063	0.16
Mesh category3	−0.053	0.948	0.278	Mesh category3	−0.067	0.935	0.226
Mesh category4	−0.146	0.864	0.002**	Mesh category4	−0.176	0.839	0.001**
Mesh category5	−0.117	0.890	0.000***	Mesh category5	−0.125	0.882	0.000***
Mesh category6	0.511	1.666	.17e−12***	Mesh category6	0.550	1.734	.43e−11***
Mesh category7	0.072	1.075	0.249	Mesh category7	0.066	1.068	0.357

The gender of first and last authors, lay summary and funding were not significant factors for the estimated number of News mentions received in both OA status and OA type models. However, the gender of first author was significantly associated with a higher probability of being mentioned in news outlets. Interestingly, Funding was significantly associated with lower odds of being mentioned in news outlets in OA status model.

Dehdarirad and Didegah: To What Extent Does the Open Access Status of Articles Predict
Their Social Media Visibility? A Case Study of Life Sciences and Biomedicine

Art. 5, page 9 of 14

Table 5: Hurdle model results for News.

Count model	Coef.	Exp(Coef.)	Sig.	Count model	Coef.	Exp(Coef.)	Sig.
OA vs. Non-OA	0.609	1.839	<2e−16***	**Gold vs. Non-Gold**	0.148	1.160	0.185
Indiv. Collaboration	0.016	1.016	0.017*	Indiv. Collaboration	0.004	1.004	0.4
Int. Collaboration	0.164	1.178	0.016*	Int. Collaboration	0.187	1.206	0.012 *
Funding	−0.001	0.999	0.913	Funding	0.008	1.008	0.562
No. Mesh Categories	0.126	1.134	1.62e−08***	No. Mesh Categories	0.114	1.121	1.08e−05***
Lay summary	0.034	1.035	0.788	Lay summary	−0.141	0.868	0.38
F1000 score	0.156	1.169	7.51e−07***	F1000 score	0.175	1.192	1.24e−06***
SNIP	0.220	1.246	<2e−16***	SNIP	0.199	1.220	2.50e−13***
Mega journals	0.251	1.286	0.017*	Mega journals	0.112	1.119	0.435
First-gender	−0.023	0.977	0.729	First-gender	0.002	1.002	0.981
Last-gender	0.004	1.004	0.955	Last-gender	0.089	1.093	0.32
Mesh category1	−0.166	0.847	0.16	Mesh category1	−0.039	0.962	0.769
Mesh category2	−0.040	0.961	0.71	Mesh category2	0.089	1.094	0.459
Mesh category3	−0.234	0.791	0.071.	Mesh category3	−0.019	0.981	0.896
Mesh category4	−0.445	0.641	0.000***	Mesh category4	−0.241	0.786	0.104
Mesh category5	−0.178	0.837	0.032*	Mesh category5	−0.107	0.899	0.246
Mesh category6	0.231	1.259	0.195	Mesh category6	0.382	1.466	0.051.
Mesh category7	−0.051	0.951	0.758	Mesh category7	0.242	1.274	0.185
Logit model	**Coef.**	**Exp(Coef.)**	**Sig.**	**Logit model**	**Coef.**	**Exp(Coef.)**	**Sig.**
OA vs. Non-OA	0.929	2.531	<2e−16***	**Gold vs. Non-Gold**	0.094	1.099	0.078.
Indiv. Collaboration	0.027	1.027	<2e−16***	Indiv. Collaboration	0.025	1.025	<2e−16***
Int. Collaboration	0.319	1.376	<2e−16***	Int. Collaboration	0.324	1.383	<2e−16***
Funding	−0.012	0.988	0.027*	Funding	−0.013	0.987	0.028
No. Mesh Categories	0.085	1.088	1.07e−14***	No. Mesh Categories	0.074	1.077	.68e−09***
Lay summary	0.695	2.005	<2e−16***	Lay summary	0.652	1.919	<2e−16***
F1000 score	0.314	1.368	<2e−16***	F1000 score	0.301	1.352	<2e−16***
SNIP	0.532	1.702	<2e−16***	SNIP	0.566	1.761	<2e−16***
Mega journals	0.036	1.037	0.483	Mega journals	−0.010	0.990	0.883
First-gender	0.121	1.128	0.000***	First-gender	0.113	1.120	0.003**
Last-gender	−0.024	0.977	0.555	Last-gender	−0.042	0.959	0.358
Mesh category1	−0.129	0.879	0.033*	Mesh category1	−0.125	0.883	0.063.
Mesh category2	−0.069	0.933	0.208	Mesh category2	−0.111	0.895	0.073.
Mesh category3	−0.074	0.929	0.269	Mesh category3	−0.046	0.955	0.537
Mesh category4	−0.170	0.844	0.011*	Mesh category4	−0.217	0.805	0.004**
Mesh category5	−0.124	0.883	0.003**	Mesh category5	−0.119	0.888	0.012*
Mesh category6	0.596	1.815	1.19e−10***	Mesh category6	0.665	1.944	.09e−11***
Mesh category7	0.073	1.076	0.387	Mesh category7	0.119	1.127	0.202

Blogs

The results of count regression models (See **Table 6**) showed that OA articles on average had significantly a higher number of blog mentions. However, the OA type of articles was not a significant factor for this platform. A unit change in the OA status increased the estimated number of blogs mentions for articles by 48.4%. The logit models showed that both the OA status and OA type significantly associated with higher probability for a paper to be mentioned in blogs.

Table 6: Hurdle model results for blogs.

Count model	Coef.	Exp(Coef.)	Sig.	Count model	Coef.	Exp(Coef.)	Sig.
OA vs. Non-OA	0.395	1.484	2.22e−06***	**Gold vs. Non-Gold**	−0.004	0.996	0.971
Indiv. Collaboration	0.013	1.013	0.016*	Indiv. Collaboration	0.008	1.008	0.16
Int. Collaboration	0.139	1.149	0.027*	Int. Collaboration	0.147	1.158	0.034*
Funding	−0.001	0.999	0.927	Funding	0.001	1.001	0.962
No. Mesh Categories	0.068	1.070	0.000***	No. Mesh Categories	0.074	1.077	0.000***
Lay summary	0.320	1.378	0.002**	Lay summary	0.343	1.409	0.009**
F 1000 score	0.179	1.196	4.48e−10***	F 1000 score	0.186	1.204	5.14e−09***
SNIP	0.259	1.296	<2e−16***	SNIP	0.274	1.315	<2e−16***
Mega journals	0.353	1.424	0.000***	Mega journals	0.370	1.448	0.004**
First-gender	−0.090	0.914	0.181	First-gender	−0.089	0.915	0.231
Last-gender	0.112	1.118	0.141	Last-gender	0.157	1.170	0.057.
Mesh category1	−0.365	0.694	0.001**	Mesh category1	−0.367	0.693	0.003**
Mesh category2	−0.327	0.721	0.001**	Mesh category2	−0.432	0.649	0.000***
Mesh category3	−0.589	0.555	6.98e−05***	Mesh category3	−0.558	0.572	0.000***
Mesh category4	−0.308	0.735	0.020*	Mesh category4	−0.320	0.726	0.029*
Mesh category5	−0.217	0.805	0.004**	Mesh category5	−0.261	0.770	0.001**
Mesh category6	0.472	1.603	0.004**	Mesh category6	0.252	1.287	0.156
Mesh category7	0.229	1.258	0.11	Mesh category7	0.303	1.354	0.058.
Logit model	**Coef.**	**Exp(Coef.)**	**Sig.**	**Logit model**	**Coef.**	**Exp(Coef.)**	**Sig.**
OA vs. Non-OA	1.005	2.731	<2e−16***	**Gold vs. Non-Gold**	0.140	1.150	0.006**
Indiv. Collaboration	0.017	1.017	9.67e−11***	Indiv. Collaboration	0.018	1.018	.26e−10***
Int. Collaboration	0.235	1.265	2.68e−13***	Int. Collaboration	0.234	1.263	.70e−11***
Funding	−0.002	0.998	0.706	Funding	−0.002	0.998	0.742
No. Mesh Categories	0.029	1.029	0.006	No. Mesh Categories	0.017	1.017	0.16
Lay summary	1.172	3.228	<2e−16***	Lay summary	1.083	2.954	<2e−16***
F 1000 score	0.302	1.352	<2e−16***	F 1000 score	0.288	1.334	<2e−16***
SNIP	0.534	1.707	<2e−16***	SNIP	0.548	1.729	<2e−16***
Mega journals	0.205	1.228	1.91e−05***	Mega journals	0.104	1.110	0.096.
First-gender	−0.070	0.933	0.041*	First-gender	−0.087	0.917	0.021*
Last-gender	−0.057	0.945	0.149	Last-gender	−0.052	0.950	0.232
Mesh category1	−0.217	0.805	0.000***	Mesh category1	−0.215	0.807	0.001**
Mesh category2	−0.085	0.918	0.103	Mesh category2	−0.055	0.946	0.338
Mesh category3	−0.479	0.619	1.17e−10***	Mesh category3	−0.404	0.668	.81e−07***
Mesh category4	−0.377	0.686	4.19e−08***	Mesh category4	−0.366	0.694	.58e−06***
Mesh category5	−0.047	0.954	0.231	Mesh category5	−0.040	0.961	0.364
Mesh category6	0.503	1.654	3.22e−08***	Mesh category6	0.577	1.781	.75e−09***
Mesh category7	0.197	1.218	0.011*	Mesh category7	0.181	1.198	0.039*

As for the controlled factors, international collaboration, journal impact (SNIP), mega journals, number of MESH topics, lay summary and F 1000 score significantly contributed to a higher probability for articles to be mentioned in blogs, both when they were examined together with the OA status and OA type factors. Interestingly, the association of lay summary with blog mentions was highly significant for both OA status and OA type regression models (both count and logit models). In the OA status model, a unit change in the lay summary (that is changing from not

Dehdarirad and Didegah: To What Extent Does the Open Access Status of Articles Predict
Their Social Media Visibility? A Case Study of Life Sciences and Biomedicine

Art. 5, page 11 of 14

having a lay abstract in the article to having a lay abstract) may increase the number of blogs mentions for articles by 37.8% which was higher than the contribution of the international collaboration (14.9%) and journal impact (29.6%).

Regarding MESH topics, the results from both logit and count models (OA vs non-OA; Gold-vs non-Gold) showed that specific MESH topics ('Anatomy', 'Diseases' and 'Chemical and drugs') had a disadvantage in terms of odds of being mentioned in blogs, as well as the estimated number of received blog mentions. The exception to these was the 'Psychiatry and Psychology' MESH category. According to the count and logit models results for OA status, the articles indexed under this MESH category were 65.4% more likely to be mentioned by blogs and to receive 60.3% more blog mentions than the rest of the articles. Furthermore, according to the logit model for OA type, articles that were categorized under the same MESH group, were also 78.1% more likely to be mentioned by blogs. The gender of last authors and funding were not significant factors for the odds of being mentioned in blogs or the estimated number of received blog mentions. However, the gender of first author was weakly associated with the lower odds of being mentioned in blogs for both OA status and OA type models.

Discussion and conclusion

This study aimed to determine whether, and to what extent, the OA status (whether an article is open or closed) and the OA type (whether an article is Gold or non-Gold) of an article in Life Sciences and Biomedicine can predict its social media visibility, when controlling for a number of important factors. These factors were individual collaboration, research funding, number of MESH topics, topic, gender of first and last authors, being a mega journal, international collaboration, lay summary, and F1000 Score.

The findings of our study revealed that the percentage of Life Sciences and Biomedicine OA articles (around 57%) mentioned on social media platforms was significantly higher than that of non-OA articles (around 36%).

Regarding the OA status, our findings showed that being open was significantly associated with a higher probability of a paper being mentioned on the social media platforms studied. Furthermore, open access articles had a higher average number of mentions on the studied platforms. This may provide insight for the European Commission as to whether they should include altmetric indicators as potential metrics for monitoring open science advancement. The highest association between OA status and the estimated number of received mentions was for Twitter (with a likelihood of 92.7% increase in the average number of tweets), whilst the lowest association was for Facebook (with a likelihood of 25.7% increase in the average of Facebook posts). The former finding is in accordance with a large-scale study by Fraser et al. (2019) who found that articles in Medical and Health Sciences received more tweets overall. This shows the importance of making an article open, regardless of type, as this makes it easier for Twitter users to access the full text of articles.

Regarding OA types (studied as Gold vs non-Gold), our findings showed that although Gold OA was the most common OA type in our studied sample, Gold OA was only associated with a higher average of Tweets counts received and a higher probability of a paper being mentioned in Tweets and blogs. While Plan S (and in that Gold OA publishing) is meant to broaden the access and subsequently increase visibility, our findings seem to suggest that social media visibility has not yet been fully reached via Gold OA in the area of Life Sciences and Biomedicine. Our findings about Twitter may be due to the fact that Twitter is a real-time microblog network. Consequently, an article might be tweeted within few hours after publication (Yu et al. 2017). Furthermore, Gold OA publications might be immediately available through preprint repositories, before the official release. This is because policies regarding deposit location, license, and embargo requirements of Gold OA might be less restrictive in comparison to other types of OA. Embargos on Green OA is an example of an access barrier (Laakso 2014). Our findings contradicted that of Holmberg et al. (2020), who found that Gold OA publications had no Twitter advantage for Finnish publications within Medicine and Health Sciences. The difference may result from the different samples studied, especially as Holmberg et al. (2020) investigated a sample of Finnish articles indexed in a national database.

With regard to the covariates for the OA status and OA type models, the results showed that overall, some covariates, such as international collaboration, journal impact and F1000 score were significantly associated with the higher probability of being mentioned on the studied platforms, as well as the estimated number of received tweets, Facebook posts, news posts and blog posts. Additionally, the 'Psychiatry and Psychology' MESH category topic was significantly associated with higher odds of visibility for both OA and Gold OA articles in all social media platforms studied. This finding is in line with Holmberg's et al. (2020) study which found that articles in the field of psychology had a clear OA advantage on Twitter. Our findings regarding news visibility of this MESH topic is in line with Kousha and Thelwall's (2019) study, which found that psychology and psychiatry were among the most frequently cited subject categories in UK newspapers. One reason for our finding, as suggested by Kousha and Thelwall (2019), might be that news outlets prefer to report research findings such as public or mental health issues that have more general interest or benefit to the public. Furthermore, the less restricted access to OA articles in comparison with non-OA articles may make it easier and faster for them to be shared and disseminated.

Our findings also showed that some factors had a higher association, an OA (dis) advantage, or were significant only on certain platforms. Some main examples of these findings are as follows. Lay summary had a high association with the number of received blog mentions (37.8%). In 2013, it was stipulated that all the National Institute for Health Research (NIHR) funded projects required lay summaries. The aim was to ensure that the results of scientific

Art. 5, page 12 of 14

Dehdarirad and Didegah: To What Extent Does the Open Access Status of Articles Predict
Their Social Media Visibility? A Case Study of Life Sciences and Biomedicine

information is disseminated in an accessible and understandable way to the general public (Kirkpatrick, et al. 2017). This is interesting, as with an increase in open access to biomedical research, some scientists and research organizations might be more likely to blog about their research. As another example, our findings showed that while the gender (female) of first author was weakly associated with higher odds of visibility for both OA and gold OA articles in news outlets, it had a weak negative association with the odds of visibility in blogs. Our findings regarding blogs is in line with Paul-Hus' et al. (2015), who found that overall, male first-authored papers had a slightly higher mean number of blogs mentions in different studied disciplines. Our findings regarding news are in contrast with Sotudeh, Dehdarirad, and Freer (2018), who found no difference between female and male first-authored papers in the field of neurosurgery in terms of visibility in news.

Regarding research funding as another example, our findings for OA status models showed that funding had a weak negative association with i) the odds of a paper being mentioned in news outlets and on Twitter ii) the average number of tweets and Facebook post counts. The later finding is in contrast with Didegah et al. (2018) which found that funding had a positive association with average number of tweets and Facebook counts received. However, our findings regarding Twitter are in line with Álvarez-Bornstein and Costas (2018) which found that biology was amongst the subject categories with higher funding rates and lower proportion of Twitter mentions.

Collectively, this study provides initial exploratory findings regarding the association between OA (status, type) and social media visibility, where controlling for several important factors. The extent to which the OA factor associates with social media counts may vary by adding or removing factors from the models. However, the current model attempted to control for several important factors. By doing so, we were able to increase the probability of obtaining a more precise and reliable association between the OA factors and the average number of received social media mentions. It is important to consider that our analysis was limited to a sample of data in the area of Life sciences and Biomedicine. Thus, the results obtained in this article are not comprehensive and readers should exercise caution with generalization of the results beyond the case studied.

Competing Interests
The authors have no competing interests to declare.

References
Adie, E. (2014). Attention! A study of open access vs non-open access articles. https://www.altmetric.com/blog/attentionoa

Alhoori, H., Ray Choudhury, S., Kanan, T., Fox, E., Furuta, R., & Giles, C. L. (2015). *On the Relationship between Open Access and Altmetrics*. Paper presented at the iConference 2015. http://hdl.handle.net/2142/73451

Álvarez-Bornstein, B., & Costas, R. (2018). Exploring the relationship between research funding and social media: disciplinary analysis of the distribution of funding acknowledgements and Twitter mention in scientific publications. *Paper presented at the 23rd International Conference on Science and Technology Indicators*, Leiden, Netherlands.

Andersen, J.P., & Haustein, S. (2015). Influence of study type on Twitter activity for medical research papers. arXiv: 1507.00154.

Chan, L., Cuplinskas, D., Eisen, M., Friend, F., Genova, Y., Guédon, J. C., Hagemann, M., Harnad, S., Johnson, R., Kupryte, R., La Manna, M., Rév, I., Segbert, M., de Souza, S., Suber, P., & Velterop, J. (2002). Budapest open access initiative. Retrieved from https://www.budapestopenaccessinitiative.org/read

Cintra, P. R., Furnival, A. C., & Milanez, D. H. (2018). The impact of open access citation and social media on leading top Information Science journals. *Investigación Bibliotecológica: Archivonomía, Bibliotecología E Información, 32*(77), 117–132. DOI: https://doi.org/10.22201/iibi.24488321xe.2018.77.57874

Didegah, F. (2014). Factors associating with the future citation impact of published articles: A statistical modeling approach. PhD thesis. University of Wolverhampton, UK.

Didegah, F., Alperin, J. P., & Haustein, S. (2018). Are Lay Summaries Useful for the Diffusion of Research on Social Platforms? *SIG/MET, ASIS&T 2018 Annual Meeting*. https://www.asist.org/SIG/SIGMET/workshop

Didegah, F., Bowman, T. D., & Holmberg, K. (2018). On the differences between citations and altmetrics: An investigation of factors driving altmetrics versus citations for Finnish articles. *Journal of the Association for Information Science and Technology, 69*(6), 832–843. DOI: https://doi.org/10.1002/asi.23934

Didegah, F., & Thelwall, M. (2013). Which factors help authors produce the highest impact research? Collaboration, journal and document properties. *Journal of informetrics, 7*(4), 861–873. DOI: https://doi.org/10.1016/j.joi.2013.08.006

European Commission. (2019). 'Plan S' and 'cOAlition S' – Accelerating the transition to full and immediate Open Access to scientific publications. Retrieved from https://ec.europa.eu/commission/commissioners/2014-2019/moedas/announcements/plan-s-and-coalition-s-accelerating-transition-full-and-immediate-open-access-scientific_en

Eysenbach, G. (2006). Citation Advantage of Open Access Articles. *PLOS Biology, 4*(5), e157. DOI: https://doi.org/10.1371/journal.pbio.0040157

Faculty Opinions. (2020). About faculty opinions. Retrieved from https://facultyopinions.com/prime/faq

Fraser, N., Momeni, F., Mayr, P., & Peters, I. (2019). Altmetrics and Open Access: Exploring Drivers and Effects. *Paper presented at The 2019 Altmetrics Workshop*, Stirling, Scotland. http://altmetrics.org/altmetrics19

Dehdarirad and Didegah: To What Extent Does the Open Access Status of Articles Predict
Their Social Media Visibility? A Case Study of Life Sciences and Biomedicine

Art. 5, page 13 of 14

Gargouri, Y., Hajjem, C., Larivière, V., Gingras, Y., Carr, L., Brody, T., & **Harnad, S.** (2010). Self-Selected or Mandated, Open Access Increases Citation Impact for Higher Quality Research. *Plos One, 5*(10), e13636. DOI: https://doi.org/10.1371/journal.pone.0013636

Hajjem, C., Harnad, S., & **Gingras, Y.** (2006). Ten-Year Cross-Disciplinary Comparison of the Growth of Open Access and How it Increases Research Citation Impact. *IEEE Data Eng. Bull., 28*, 39–46.

Haustein, S., Bowman, T. D., & **Costas, R.** (2015). Interpreting "altmetrics": Viewing acts on social media through the lens of citation and social theories. arXiv:1502.05701. DOI: https://doi.org/10.1515/9783110308464-022

Hersh, G., & **Plume, A.** (2016). Citation metrics and open access: what do we know? *Elsevier Connect.* Retrieved from https://www.elsevier.com/connect/citation-metrics-and-open-access-what-do-we-know

Hilbe, J. M. (2011). Negative Binomial Regression, second edition (5th print, 2012). Cambridge, UK: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511973420

Holmberg, K., Hedman, J., Bowman, T. D., Didegah, F., & **Laakso, M.** (2020). Do articles in open access journals have more frequent altmetric activity than articles in subscription-based journals? An investigation of the research output of Finnish universities. *Scientometrics, 122*(1), 645–659. DOI: https://doi.org/10.1007/s11192-019-03301-x

Hua, F., Sun, H., Walsh, T., Worthington, H., & **Glenny, A.-M.** (2016). Open access to journal articles in dentistry: Prevalence and citation impact. *Journal of Dentistry, 47*, 41–48. DOI: https://doi.org/10.1016/j.jdent.2016.02.005

Kanjilal, U., & **Das, A. K.** (2015). *Introduction to open access.* Paris, France: United Nations Educational, Scientific and Cultural Organization.

Kirkpatrick, E., Gaisford, W., Williams, E., Brindley, E., Tembo, D., & **Wright, D.** (2017). Understanding Plain English summaries. A comparison of two approaches to improve the quality of Plain English summaries in research reports. *Research Involvement and Engagement, 3*(1), 17. DOI: https://doi.org/10.1186/s40900-017-0064-0

Kousha, K., & **Thelwall, M.** (2019). An Automatic Method to Identify Citations to Journals in News Stories: A Case Study of UK Newspapers Citing Web of Science Journals. *Journal of Data and Information Science, 4*(3), 73–95. DOI: https://doi.org/10.2478/jdis-2019-0016

Kriegeskorte, N., Walther, A., & **Deca, D.** (2012). An emerging consensus for open evaluation: 18 visions for the future of scientific publishing. *Frontiers in computational neuroscience, 6*, 94. DOI: https://doi.org/10.3389/fncom.2012.00094

Laakso, M. (2014). Green open access policies of scholarly journal publishers: a study of what, when, and where self-archiving is allowed. *Scientometrics, 99*(2), 475–494. DOI: https://doi.org/10.1007/s11192-013-1205-3

McCabe, M. J., & **Snyder, C. M.** (2015). Does Online Availability Increase Citations? Theory and Evidence from a Panel of Economics and Business Journals. *The Review of Economics and Statistics, 97*(1), 144–165. DOI: https://doi.org/10.1162/REST_a_00437

McKiernan, E. C., Bourne, P. E., Brown, C. T., Buck, S., Kenall, A., Lin, J., McDougall, D., Nosek, B. A., Ram, K., Soderberg, C. K., & **Spies, J. R.** (2016). How open science helps researchers succeed. *eLife, 5*, e16800. DOI: https://doi.org/10.7554/eLife.16800

Paul-Hus, A., Sugimoto, C., Haustein, S., & **Larivière, V.** (2015). *Is there a gender gap in social media metrics?', in ISSI 2015-15th International conference of the International Society for Scientometrics and Informetrics.* Paper presented at the ISSI 2015-15th International conference of the International Society for Scientometrics and Informetrics, Instanbul, Turkey.

Piwowar, H., Priem, J., Larivière, V., Alperin, J. P., Matthias, L., Norlander, B., Farley, A., West, J., & **Haustein, S.** (2018). The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles. *PeerJ, 6*, e4375. DOI: https://doi.org/10.7717/peerj.4375

Santamaría, L., & **Mihaljević, H.** (2018). Comparison and benchmark of name-to-gender inference services. *PeerJ Computer Science, 4*, e156. DOI: https://doi.org/10.7717/peerj-cs.156

Sotudeh, H., Dehdarirad, T., & **Freer, J.** (2018). Gender differences in scientific productivity and visibility in core neurosurgery journals: Citations and social media metrics. *Research Evaluation, 27*(3), 262–269. DOI: https://doi.org/10.1093/reseval/rvy003

Spezi, V., Wakeling, S., Pinfield, S., Creaser, C., Fry, J., & **Willett, P.** (2017). "Open-access mega-journals: The future of scholarly communication or academic dumping ground? A review". *Journal of Documentation, 73*(2), 263–283. DOI: https://doi.org/10.1108/JD-06-2016-0082

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: https://doi.org/10.1371/journal.pone.0064841

Van Leeuwen, T., Robinson-Garcia, N., & **Costas, R.** (2019). The Matthew Effect of Plan S: Is Gold OA Publishing Mainly a Business Model Fitting the Rich in Science? *Paper presented at the The 6th conference on scholarly publishing in the context of open science*, Zadar, Croatia. http://pubmet.unizd.hr/pubmet2019/talks/the-matthew-effect-of-plan-s-is-gold-oa-publishing-mainly-a-business-model-fitting-the-rich-in-science

Wang, X., Liu, C., Mao, W., & **Fang, Z.** (2015). The open access advantage considering citation, article usage and social media attention. *Scientometrics, 103*(2), 555–564. DOI: https://doi.org/10.1007/s11192-015-1547-0

Yu, H., Xu, S., Xiao, T., Hemminger, B. M., & **Yang, S.** (2017). Global science discussed in local altmetrics: Weibo and its comparison with Twitter. *Journal of Informetrics, 11*(2), 466–482. DOI: https://doi.org/10.1016/j.joi.2017.02.011