

A influência do Acesso Aberto nas citações e nas métricas alternativas: o caso do ISCTE-IUL

António Luís Lopes

PhD, ISCTE-Instituto Universitário de Lisboa (ISCTE-IUL), Lisboa, Portugal.

Investigador, ISCTE-Instituto Universitário de Lisboa (ISCTE-IUL), Lisboa, Portugal.

<https://ciencia.iscte-iul.pt/authors/antonio-luis-lopes>

E-mail: alsl@iscte-iul.pt

Data de submissão: 02/09/2019. Data de aprovação no ConfOA: 12/06/2019. Data de publicação

RESUMO

A bandeira da ciência aberta assenta no princípio de toda e qualquer investigação científica estar disponível e acessível ao mundo inteiro, independentemente da condição social ou económica do indivíduo a que ela pretende aceder. Nesse sentido, e intuitivamente, seria de esperar que o facto de determinada produção científica estar em acesso aberto contribuiria para aumentar o seu impacto, nomeadamente ao nível do número de citações que viesse a obter. Os estudos realizados a esse nível não chegam a um consenso, sendo que alguns apontam o acesso aberto como um fator que influencia o número de citações e havendo outros que dizem especificamente o contrário. Nesta publicação é exposto o caso do ISCTE-IUL, universidade pública em Lisboa, Portugal, onde é testada a influência do acesso aberto no impacto da divulgação das produções científicas dos investigadores. Conclui-se que, apesar da influência do acesso aberto no número de citações existir, ela é muito ténue em face de outros fatores como o impacto da própria revista onde se publica. Além disso, conclui-se também que a influência do acesso aberto é mais evidente nas métricas alternativas, como as referências em redes sociais que um artigo tem.

Palavras-chave: Acesso Aberto. Citações. Métricas alternativas. Avaliação.

The influence of open access in citations and alternative metrics: the case of ISCTE-IUL

ABSTRACT

The flag of open science rests on the principle that each and every scientific research should be available and accessible to the whole world, independently of the social and economic status of the individual that is trying to access it. In this sense, and intuitively, one would hope that the fact that a specific scientific production is openly available and accessible would contribute to increase the impact of that production, namely in the number of citations that it would amass. Studies on this matter are unable to reach a consensus, with some indicating open access as a factor that influences the number of citations and others that explicitly state otherwise. In this paper, we present the case of ISCTE-IUL, a public University in Lisboa, Portugal, where we studied the influence of open access on the impact of scientific productions of our researchers. We have concluded that, on one hand, even though the influence of open access on the number of citations exists, this connection is quite tenuous compared to other factors like the impact of the journal itself where the publication is published. On the other hand, we have concluded that the influence of open access manifests more clearly in the alternative metrics, like the number of references a publication has in social networks.

Keywords: Open Access. Citations. Alternative Metrics. Evaluation.

Título La influencia del acceso abierto en las citas y las métricas alternativas: el caso de ISCTE-IUL

Resumen

El estandarte de la ciencia abierta se basa en el principio de que toda la investigación científica está disponible y accesible para todo el mundo, independientemente de la condición social o económica de la persona a la que busca acceder. En este sentido, e intuitivamente, se esperaría que el hecho de que una determinada producción científica fuera de acceso abierto contribuiría a aumentar su impacto, es decir, en términos del número de citas que obtendría. Los estudios a este nivel no llegan a un consenso, y algunos apuntan al acceso abierto como un factor que influye en el número de citas y otros dicen específicamente lo contrario. En este artículo exponemos el caso de ISCTE-IUL, una universidad pública en Lisboa, Portugal, donde probamos la influencia del acceso abierto en el impacto de la difusión de las producciones científicas de los investigadores. Llegamos a la conclusión de que, por un lado, a pesar de la influencia del acceso abierto en el número de citas, es muy leve en comparación con otros factores, como el impacto de la propia revista. Por otro lado, también concluimos que la influencia del acceso abierto es más evidente en las métricas alternativas, como las referencias en las redes sociales que tiene un artículo.

Palabras clave: Acceso abierto. Citas. Métricas alternativas. Evaluación.

INTRODUÇÃO

O impacto de uma publicação científica é geralmente medido através do seu número de citações. Esta métrica reflete a importância que a comunidade científica atribui à publicação através das referências que são feitas noutras publicações. Por representar a imputação de valor que terceiros fazem a cada publicação, o número de citações é normalmente usado em processos de avaliação de desempenho de investigadores e instituições. Dada a importância que é atribuída a essa métrica, é relevante perceber o que condiciona a geração de citações: que fatores influenciam que uma publicação tenha mais ou menos citações?

Ao longo das últimas décadas, muitos estudos foram sendo feitos para perceber quais são os fatores que influenciam esse indicador, e a maioria aponta o impacto ou prestígio da própria revista como o principal fator. Naturalmente, esse comportamento é compreensível, uma vez que o impacto ou importância de uma revista é dado pelo número de citações que ela tem em função do número de artigos publicados. Logo, entramos num ciclo que se alimenta a si próprio: revistas com mais citações têm mais prestígio; por sua vez, os autores tendem a tentar publicar mais em revistas com prestígio; isto

atrai os artigos com maior qualidade; a revista acaba por publicar artigos que têm mais visibilidade, logo, são mais citados; o que por sua vez aumenta as próprias citações, ampliando também o prestígio da revista. E assim sucessivamente.

Com o advento do paradigma do Acesso Aberto e o seu papel na difusão de uma Ciência Aberta, é importante analisar se ele contribui de algum modo para expandir o impacto de uma publicação científica. A expectativa é a de que artigos que estão imediatamente disponíveis após publicação para consulta livre por qualquer indivíduo no mundo terão maior tendência para serem citados. Contudo, estudos realizados nesse sentido não chegam a um consenso. Tanto há estudos (CRAIG *et al.*, 2007) (DAVIS, 2011) (CHUA *et al.*, 2017) que referem que não existe uma relação causal ou de influência entre o facto de uma publicação se encontrar em acesso aberto e o número de citações que ela tem, quanto há estudos (ANTELMAN, 2004) (EYSENBACH, 2006) (WANG *et al.*, 2015) que declaram exatamente o contrário: existe de facto uma relação causal entre o acesso aberto e o número de citações que uma publicação possui.

A medição do impacto de uma publicação pode também ser feita recorrendo a um conjunto de métricas alternativas ao número de citações, nomeadamente, o número de referências que são feitas à publicação noutros meios como redes sociais na internet ou plataformas on-line de âmbito académico-científico. Apesar de tais referências realizadas em meios alternativos não garantirem rigor científico semelhante às citações realizadas em revistas científicas, representam ainda assim uma ferramenta importante de divulgação em massa que poderá, no futuro, contribuir para gerar mais citações para os artigos científicos a que concernem. E a verdade é que, mesmo nos estudos citados anteriormente que alegam não haver influência entre o Acesso Aberto e o número de citações de uma publicação, existe uma influência mais forte entre o Acesso Aberto e as métricas alternativas. É o caso de estudos como o de Davis (2011), que apesar de não encontrar uma ligação causal entre publicações em Acesso Aberto e o seu número elevado de citações, encontra uma ligação causal forte entre publicações em Acesso Aberto e número alto de visualizações, downloads e referências em redes sociais. Esse impacto alternativo na divulgação em massa da publicação pode funcionar como o catalisador que permitirá à publicação entrar na esfera de observação de outros investigadores, aumentando sua probabilidade de vir a ser citada no futuro.

METODOLOGIA

Neste artigo é apresentado um estudo de caso realizado no Instituto Universitário de Lisboa (ISCTE-IUL), universidade pública em Lisboa, Portugal. No ISCTE-IUL, todos os docentes e investigadores fazem uso do CRIS local, Ciência-IUL (LOPES, 2019), para divulgar suas produções científicas. À data (setembro de 2019), existem cerca de 35 mil produções científicas registadas no sistema, que geraram mais de 84 mil citações.

O estudo de caso baseia-se nos seguintes pressupostos:

- A informação de citações das publicações foi recolhida automaticamente a partir dos serviços Web of Science e Scopus. Para cada publicação, o número de citações considerado foi o máximo dos dois valores obtidos nas duas bases de dados de indexação e citação.
- A informação sobre quartis das revistas foi recolhida a partir da informação disponibilizada pelo Journal Citation Reports (Web of Science) e Scimago Journal Ranking (Scopus). Para cada publicação, o quartil considerado foi o mínimo encontrado na categoria da publicação nos dois serviços para o ano de referência da publicação. O quartil funciona numa escala inversa: quanto menor o quartil, maior é o impacto da revista. Foi decidido usar o quartil em vez do Impact Factor (IF) porque o quartil reflete o posicionamento do IF da revista no contexto da área em que esta se insere, logo é mais indicativo do verdadeiro impacto da revista.
- A informação sobre métricas alternativas foi obtida através dos serviços Altmetric (<https://www.altmetric.com/>), Dimensions (<https://www.dimensions.ai/>) e PlumX (<https://plumanalytics.com/learn/about-metrics/>). A API PlumX encontra-se livremente disponível para instituições que já subscrevam o serviço Scopus. Os dados recolhidos foram:
 - Referências feitas no Twitter, Facebook, Google, Mendeley, Notícias, Wikipedia e Patentes;
 - Número de cliques, downloads e visualizações.
- A informação sobre o nível de Acesso Aberto de uma publicação foi recolhida automaticamente a partir do serviço *Sherpa Romeo*, e é considerada a dois níveis:
 - o parâmetro “Romeo Color”, que classifica a revista com uma cor que designa o nível de Acesso Aberto que ela assume: *gray* (para revistas não classificadas ou em revisão); *white* (para revistas fechadas); *yellow* (permite

arquivar o pré-print); *blue* (permite arquivar o pós-print ou a versão de editora); *green* (em Acesso Aberto total);

- uma fórmula ponderada dos três parâmetros de permissão de arquivo (pré-print, pós-print, versão da editora).
- Só foram consideradas as publicações para as quais foi possível obter a informação descrita anteriormente. No total, foram consideradas 6.419 publicações para as quais foi possível obter pelo menos um dos elementos da informação descrita.

O estudo de caso realizado foi feito usando o SPSS para proceder à análise estatística através da regressão linear. Para esse efeito, foi necessário converter todos os dados recolhidos em formato numérico. Naturalmente, as citações e as referências nas métricas alternativas foram usadas tal como foram recolhidas, por já corresponderem a valores numéricos.

No que toca aos quartis das revistas (de Q1 a Q4), foi usada uma conversão direta (Q1 = 1, Q2 = 2, Q3 = 3 e Q4 = 4). Adicionalmente, para os casos em que a classificação de uma revista Q1 se enquadrava na classificação de Top 10% ou Top 5% da área científica, foi feita também uma conversão correspondente, isto é, *Top 10%* = -1 e *Top 5%* = -5. Assim, podemos diferenciar o verdadeiro impacto de revistas dentro da categoria Q1. Portanto, para este efeito, a classificação possível de publicações em termos de quartis é: [-5, -1, 1, 2, 3, 4].

No que toca à informação de Acesso Aberto, o mapeamento entre os valores do parâmetro “*Romeo Color*” e o valor numérico foi o seguinte: *Gray* (não considerada); *White* = 1; *Yellow* = 2; *Blue* = 3; *Green* = 4.

A fórmula usada para o cálculo do valor ponderado dos três parâmetros de permissão de arquivo (pré-print, pós-print, versão da editora) foi a seguinte (para cada parâmetro, o valor possível é 0 – em que indica que não é permitido – e 1 – em que indica que é permitido):

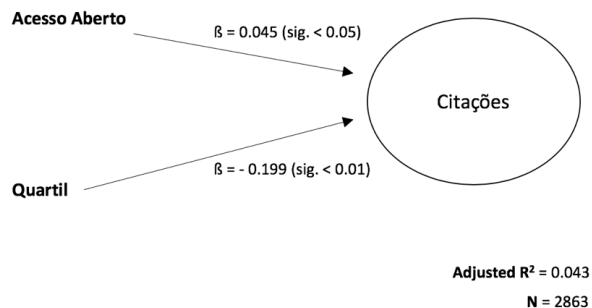
$$\text{Pré-print} + \text{pós-print} * 2 + \text{versão da editora} * 3$$

Dessa fórmula retira-se, portanto, a indicação do nível de acesso aberto que determinada publicação pode assumir (um valor entre [0 - 6]), sendo que 6 representa uma publicação que está totalmente disponibilizada em Acesso Aberto. Na análise estatística realizada, foi usado o máximo dos dois valores (Romeo Color e Acesso Aberto).

ANÁLISE E DISCUSSÃO DOS RESULTADOS

O estudo iniciou-se com a necessidade de verificar a relação que efetivamente existe entre o número de citações de uma publicação e o facto de ela estar ou não em Acesso Aberto. Ao mesmo tempo, pretende-se analisar se essa variável (estar ou não em Acesso Aberto) é influenciada pela presença da outra variável que representa o impacto efetivo da revista (dado pelo quartil).

Figura 1 – Influência do Acesso Aberto e do quartil no número de citações



A figura 1 apresenta o resultado da primeira análise de regressão com ambas as variáveis (Acesso Aberto e Quartil) e a sua contribuição para a variável Citações, para um universo de 2.863 publicações (para as quais foi possível apurar toda a informação necessária de acesso aberto, quartil e citações).

Como se pode ver¹, há de facto uma influência ténue ($\beta = 0.045$; sig. = 0.014) do Acesso Aberto no número de citações, isto é, quanto maior for o

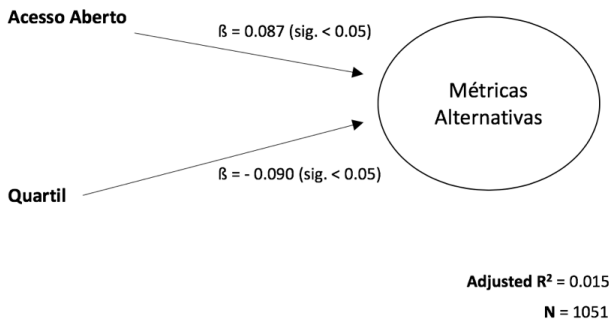
1 O que se procura nesta análise é que o valor de β seja o mais elevado possível (valor positivo se a influência é positiva, ou valor negativo se a influência é negativa) e o valor de sig. seja o mais reduzido possível.

nível de acesso aberto de uma publicação, maior é a contribuição para o aumento do número de citações. No entanto, a variável Quartil aparenta ter uma influência superior ($\beta = -0.199$; sig. = 0.000). O valor é negativo porque a influência é inversa, isto é, quanto menor o quartil, maior é a probabilidade de a publicação ter mais citações.

A importância da variável Quartil em relação à variável Acesso Aberto é também evidente no valor do *Adjusted R*². Quando se considera apenas a influência da variável Acesso Aberto, o valor do *Adjusted R*² é de apenas 0.004 ($\beta = 0.067$, sig. = 0.000). É quando se introduz a variável Quartil que esse valor passa para 0.043, evidenciando uma relação muito mais forte com essa variável.

A primeira análise reflete a importância que o prestígio da revista onde a publicação se insere (em função do seu quartil) tem no número de citações que a publicação acaba por receber. Em seguida, pretende-se averiguar se o mesmo efeito se aplica às métricas alternativas. Foi feita uma segunda análise de regressão, em que a variável dependente representa o total de referências realizadas ao nível de redes sociais e outras plataformas (Facebook, Google, Twitter, presenças em Patentes, notícias e Wikipedia).

Figura 2 – Influência do Acesso Aberto e quartil nas métricas alternativas



A figura 2 apresenta o resultado da análise de regressão com ambas as variáveis (Acesso Aberto e Quartil) e a sua contribuição para a variável Métricas Alternativas, para um universo de 1051 publicações (para as quais foi possível apurar toda a informação necessária de acesso aberto, quartil e métricas alternativas). Como se pode ver, a influência do Acesso Aberto nas Métricas Alternativas ($\beta = 0.087$; sig = 0.05) é maior do que para o Número de Citações. A variável Quartil aparenta ter influência inferior ($\beta = -0.090$; sig = 0.000) nas Métricas Alternativas do que tem no Número de Citações. A variação no valor do Adjusted R² com a introdução da variável Quartil também é inferior à análise anterior. Quando se considera apenas a influência da variável Acesso Aberto, o valor do Adjusted R² é de 0.008 ($\beta = 0.093$, sig. = 0.002). Quando se introduz a variável Quartil, esse valor passa para 0.015, evidenciando alguma importância da variável Quartil, mas não tanto como na influência evidente no número de citações.

CONCLUSÕES

Essas análises sugerem que a ligação entre o Acesso Aberto e o número de citações é muito ténue, mas a ligação com as métricas alternativas é mais forte. Isto parece ir ao encontro dos estudos citados no início, em que a influência do acesso aberto nas citações é difícil de estabelecer (sendo aparente nuns casos, e noutros não) mas onde é mais fácil estabelecer uma relação com a divulgação das publicações em plataformas como as redes sociais e outras.

Importa realçar que nas análises realizadas o Adjusted R² é de facto baixo, o que permite concluir que apesar de significativos, os modelos estatísticos apresentados consideram-se pouco explicativos. Isto leva-nos a crer que, apesar de haver uma correlação nas variáveis estudadas, há outros fatores que explicam a origem das citações (provavelmente de forma mais evidente).

O trabalho futuro incidirá sobre esses outros fatores (internos ou externos) que poderão explicar a relação com o impacto da publicação.

REFERÊNCIAS

ANTELMAN, K. Do open-access articles have a greater research impact? *College & research libraries*, [S.l.], v. 65, n.5, p.372-382, 2004.

CHUA, S.K. *et al.* T. The impact factor of an open access journal does not contribute to an article's citations. *F1000Research*, [S.l.], v. 6, 2017.

CRAIG, I. D. *et al.* Do open access articles have greater citation impact?: a critical review of the literature. *Journal of Informetrics*, [S.l.], v.1, n. 3, p.239-248, 2007.

DAVIS, P.M. Open access, readership, citations: a randomized controlled trial of scientific journal publishing. *The FASEB Journal*, [S.l.], v. 25, n. 7, p.2129-2134, 2011.

EYSENBACH, G. Citation advantage of open access articles. *PLoS biology*, [S.l.], v. 4, n. 5, p.e157, 2006.

LOPES, A.L. Integrating a local CRIS with the PTCRIS synchronization ecosystem. *Procedia Computer Science*, [S.l.], v.146, p. 166-172, 2019.

WANG, X. *et al.* The open access advantage considering citation, article usage and social media attention. *Scientometrics*, [S.l.], v. 103, n. 2, p.555-564, 2015.

AGRADECIMENTOS

O autor deseja agradecer o apoio da colega Raquel Velada do ISCTE-IUL, pelo seu contributo valioso no trabalho de análise estatística deste estudo.