

# No evidence of citation bias as a determinant of STEM gender disparities in US Biochemistry, Genetics and Molecular Biology research<sup>1</sup>

Mike Thelwall, Tamara Nevill, University of Wolverhampton, UK.

The lack of females in many Science Technology, Engineering and Mathematics (STEM) subjects in the USA is an ongoing concern, with many initiatives attempting to redress this imbalance. Some life sciences are apparently areas of relatively good practice, with higher proportions of female researchers than most other STEM subjects. This paper assesses gender differences in research contributions to 14 biochemistry, genetics or molecular biology specialisms in the USA 1996-2014/8, seeking evidence of trends in publishing and citation impact that may give insights into female success. With four exceptions (Biochemistry, Biophysics, Biotechnology, and Structural Biology), the fields achieved or maintained at least 40% female first authors by 2018, with Developmental Biology and Endocrinology both attaining female first author majorities. A regression analysis found close to gender parity overall in citation impact but a small male first author citation advantage in more fields than the opposite: an up to 3% increase in logged citation ratio to the world mean. This was partly due to males first authoring with larger teams. Fields with relatively many females did not favour female-led research with more citations, however.

**Keywords:** Gender bias; Citation analysis; Bibliometrics; Life sciences

## Introduction

Female underrepresentation in Science Technology, Engineering and Mathematics (STEM) subjects continues in the USA despite girls outperforming boys in academic grades (O'Dea, Lagisz, Jennions, & Nakagawa, 2018) and female majorities elsewhere for faculty (NCES 2019) and publishing (Elsevier, 2017; Thelwall, Bailey, Tobin, & Bradshaw, 2019). The life sciences are an important STEM exception, however (Sax, Lim, Lehman, & Lonje-Paulson, 2018). A recent analysis of gender in publishing found Biochemistry, Genetics and Molecular Biology to be the most gender equal Scopus STEM category in the US in terms of publishing 2011-2015, with 42% of publishing authors being female compared to, for example, 40% in Agricultural and Biological Sciences, and 21% in Engineering (Elsevier, 2017, calculations from Fig 1.3). Since there have been claims that under-citation of female research may hinder female careers (Guglielmi, 2019; Larivière, Ni, Gingras, Cronin, & Sugimoto, 2013), this article investigates whether gender citation differentials associate with the proportion of females within Biochemistry, Genetics and Molecular Biology. A positive result would help to partially support the conclusion that citation biases could help to explain the relative attractiveness of Biochemistry, Genetics and Molecular Biology compared to other STEM subjects to females, although other types of evidence would be needed to assess causality.

Bibliometrics may be part of the STEM problem of underrepresentation of females because career-based statistics (h-index, total citations) are gender biased by ignoring factors disproportionately affecting females, such as career gaps and part-time working for carer responsibilities (Hopkins, Jawitz, McCarty, Goldman, & Basu, 2013). There are no

---

<sup>1</sup> Thelwall, M. & Nevill, T. (in press). No evidence of citation bias as a determinant of STEM gender disparities in US Biochemistry, Genetics and Molecular Biology research. *Scientometrics*.

large-scale detailed studies of gender for Biochemistry, Genetics and Molecular Biology, but an evaluation of Science, Technology, Engineering, Mathematics and Medicine (STEMM) publishing in PubMed and arXiv found a slowly closing gender gap that was largest in wealthy countries (Holman, Stuart-Fox, & Hauser, 2018). Within the Biochemistry, Genetics and Molecular Biology category, it estimated the 2016 global female proportion of first authors to be between 25% (Bioinformatics) and 45% (Biotechnology). In India, solo male authorship is far more common than solo female authorship in the life sciences and males dominate overall (Garg & Kumar, 2014), showing that gender patterns vary internationally (see also: Elsevier: 2017).

This paper assesses gender differences in contributions to, and citation impact within, all Biochemistry, Genetics and Molecular Biology narrow research fields, following Scopus All Science Journal Classification codes ([www.elsevier.com/solutions/scopus/how-scopus-works/content](http://www.elsevier.com/solutions/scopus/how-scopus-works/content)).

## Data and Methods

Records for all US-authored journal articles classified in at least one of the 14 Biochemistry, Genetics and Molecular Biology specialisms in Scopus were obtained between November and December 2018 (for 1998-2017) or the end of January 2019 (for 2018 articles). The analysis is for articles with US authors only to avoid contamination by international differences in research impact and research collaboration effects (Gazni & Didegah, 2011). The US has been the major single contributor to this Scopus category. Articles published 2015-18 were excluded from the citation analysis component because several years are needed for citation counts to stabilise and shrink within-year citation differentials (de Araújo, de Oliveira, de Oliveira Brito, et al., 2012). First and last author genders were detected using a list of gendered first names that has an error rate of 1.5% on US academics but leaves some genders unknown and overestimates the proportion of females by 1% (based on comparisons with human-checked data in: Thelwall et al., 2019). Articles with either the first or last author gender unknown were removed. The first and last author genders were detected for both the first and last author in most articles in nearly all fields, with Genetics being the lowest (49% gendered).

The first and last authors are normally the main contributors to a paper (Larivière, Desrochers, Macaluso, Mongeon, Paul-Hus, & Sugimoto, 2016), with the last author sometimes being the laboratory head, the person that won the funding, or another senior figure that may have guided the topic choice and research design (Tscharrntke, Hochberg, Rand, Resh, & Krauss, 2007). Exceptions to this rule add noise to the data and tend to reduce the size of any differences. Thus, the underlying differences may be slightly larger than apparent from the data. Larger team size (number of authors) associates with higher citation impact (Larivière, Gingras, Sugimoto, & Tsou, 2015) so this factor was included within the analysis.

Average citation counts should not be compared between sets of articles for multiple reasons. They are highly skewed, so the arithmetic mean is inappropriate (Zitt, 2012). They increase over time and so should not be compared between years. Average citation rates vary between specialisms (Waltman, van Eck, van Leeuwen, Visser, & van Raan, 2011), including for multidisciplinary research (Levitt & Thelwall, 2008) and so should not be compared between fields or between monodisciplinary and multidisciplinary research. To address these issues, each citation count was replaced with the (field and year) Normalised Log-transformed Citation Score (NLCS) before analysis (Thelwall, 2017). The log

transformation is  $\ln(1 + c)$ , reduces skewing in sets of articles without removing uncited articles. The field normalisation is  $NLCS = \frac{\ln(1+c)}{\ln(1+c)}$ , where the denominator is the average of the averaged  $\ln(1 + c)$  logged citation counts for each field in which the article is classified. NLCS can be averaged to give the Mean NLCS (MNLCS) (Thelwall, 2017), which can be compared between sets of articles from different fields and years without bias, by design. This is because each NLCS is a ratio to the world average for the field(s) and year of publication. MNLCS confidence intervals can safely be calculated for sets of NLCS using the standard normal (t-distribution) formulae even though they are not normally distributed, if the skewness and excess kurtosis are moderate (below 3).

Since average citation rates have varied over time in the USA, a simple average of male and female NLCS could be unfair due to changing gender proportions over time and would have an unnecessarily high variance. Instead, linear regression was used with NLCS as the dependant variable, first and last author genders and team size (1, 2, 3, 4, 5+) as independent binary variables, and publication year as a set of binary variables (one for each year). This is preferable to multilevel modelling to deal with the irrelevant publication year effect because there is sufficient data to avoid the necessity of making assumptions about the distribution of NLCS variations by year. Thus, the main regression equation incorporates first author  $F$  (the main independent variable of interest) and last author  $L$  gender, variables  $A_2$  to  $A_5$  for 2 to 5+ authors, and variables  $P_y$  for each publication year from 1997 to 2014; all are binary variables.

$$NLCS = \alpha + \beta_1 F + \beta_2 A_2 + \beta_3 A_3 + \beta_4 A_4 + \beta_5 A_{5+} + \beta_6 L + \beta_7 P_{97} + \dots + \beta_{24} P_{14}$$

The residuals from an Ordinary Least Squares (OLS) regression showed heteroscedasticity, making the standard error estimates unsafe. This was due to different spreads of citation counts between years, also inducing different spreads between genders. Thus, a weighted least squares algorithm was used instead, which is safe in the presence of heteroscedasticity (Yohai, 1987). Outliers were checked for in the residuals, but no problems were found.

For model triangulation, simplified regression models were also run without the team size variables and without the last author gender variable. Removing a variable equates to assuming that it is under the control of the remaining regression variables (e.g., team size under control of the first author, rather than being independent).

## Results

By 2018, 10 out of 14 fields had at least 40% female first-authored articles (scope: domestic US articles only), with the exceptions being Biophysics (29%), Biotechnology (32%), Structural Biology (32%) and Biochemistry (36%) (Table 1). Only two fields had female majorities: Developmental Biology (52%) and Endocrinology (57%). All fields had also experienced an increase in the proportion of female first-authored articles 1996-2018, varying from 1% (Aging) to 20% (Physiology). This increase has been steady in all cases (Online supplement Figures S1-S13) except Structural Biology, where variations between years and the lack of a pattern combine to make an overall gender change trend uncertain.

Female last authors have also experienced an increase in proportions for all fields 1996-2018 (Table 1) but this proportion is much lower than that of first authors, presumably because a higher proportion of senior academics are male and these are more likely to occupy last place in author lists. Even in Aging, which had 44-45% female authors 1996-2018, the proportion of female last authors is low (38%).

Table 1. Gender identification and share statistics for 14 Biochemistry, Genetics and Molecular Biology Scopus narrow fields 1996-2018. Fields are in descending order of female first author proportions in 2018.

Narrow field	Articles with US only authors	Gendered first & last authors %	Female first in 1996 %	Female first in 2018 %	Change 1996-2018 %	Female last in 1996 %	Female last in 2018 %	Change 1996-2018 %
Endocrinology	46902	54.1	37.7	56.8	19.0	22.7	41.2	18.5
Developmental Biology	39905	58.3	39.7	51.6	11.9	20.7	37.7	17.0
Aging	9305	63.6	44.1	45.3	1.2	35.5	37.2	1.7
Cancer Research	101918	54.0	30.3	45.1	14.8	18.5	33.3	14.8
Physiology	92824	56.2	24.6	44.9	20.3	17.0	28.5	11.5
Genetics	141184	49.0	31.7	44.8	13.1	19.7	31.8	12.1
Clinical Biochemistry	45725	53.4	27.9	44.0	16.1	21.0	30.5	9.6
Cell Biology	164124	50.1	34.9	43.4	8.5	18.9	27.7	8.8
Molecular Biology	239591	50.9	33.4	41.9	8.5	18.0	28.0	10.0
Molecular Medicine	60070	51.5	22.5	40.3	17.8	15.3	24.7	9.4
Biochemistry	266485	50.0	28.8	35.9	7.0	15.1	23.2	8.2
Structural Biology	22725	52.7	26.1	32.4	6.3	16.5	21.1	4.7
Biotechnology	61364	45.6	25.9	32.2	6.3	14.4	22.9	8.6
Biophysics	62518	49.3	24.1	28.6	4.5	13.8	20.3	6.4

First and last author gender were compared to team size for background information, using data from the most recent year. Data cannot be aggregated between years because older years had smaller teams and fewer females. In most fields (11/14), female first authors were in teams in 2018 that were larger than those of male first authors, with the difference being statistically significant in four cases (Table 2). The converse is true for female last authors (5/14, two statistically significant), presumably due to senior male team leaders often being last authors within larger teams.

Table 2. Average (geometric mean) team sizes in 2018 for Scopus narrow Biochemistry, Genetics and Molecular Biology fields (US only) by first and last author gender. Fields are in descending order of female first author proportions in 2018.

Narrow field	Articles in 2018	Female first team size	Male first team size	Female last team size	Male last team size
Endocrinology	962	5.14	5.12	5.05	5.19
Developmental Biology	841	4.88	4.41	4.60	4.67
Aging	296	5.42	5.49	5.48	5.45
Cancer Research	2638	7.32	7.07	7.00	7.27*
Physiology	2012	4.52	4.30	4.29	4.44
Genetics	3201	5.05***	4.46	4.67	4.74
Clinical Biochemistry	727	5.12	5.15	4.85	5.27
Cell Biology	2514	5.16	4.94	4.65	5.19***
Molecular Biology	4360	5.09***	4.64	4.84	4.82
Molecular Medicine	1383	5.84	6.05	5.87	5.99
Biochemistry	3550	4.78***	4.29	4.52	4.44
Structural Biology	445	4.09	4.05	4.39	3.98
Biotechnology	1392	4.44**	3.96	3.93	4.16
Biophysics	1209	4.24	4.03	4.22	4.05

\*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$  Wilcoxon male vs. female tests; the larger of the two values is starred

In the main regression models, there are statistically significant first or last author gender effects in almost all (13/14) fields (Table 3) and all fields had a gender effect in at least one of the four models tested. The effect sizes are always small, and much smaller than the influence of team size (Table 3, 4). In the main model, the small statistically significant effects are negative for first author females in nine fields and positive in two. For last author females, the effect is negative in seven fields and positive in one. The largest gender effect is only 3.4%: a 3.4% increase in the average ratio of logged citations to the world mean.

When team size is removed from the model (i.e., modelled as an output of the work of the first and/or last author rather than regarded as an input), the last author female disadvantages largely remain but the first author effects shrink substantially (2/14 positive, 3/14 negative in Model 2; 2/14 positive, 3/14 negative in Model 3) (Table 4). Thus, part of the male citation advantage in some fields associates with first authoring research by larger teams.

Although team size has a more substantial association than gender with citation impact in all fields, its contribution varies considerably between fields. Whilst a 5-authored Biotechnology article can expect a 63% higher ratio of logged citations compared to the world average than a solo-authored paper, the figure for Structural Biology is only 24% (Table 3). Thus, whilst team science seems to generate more citation impact in all fields, the effect is most substantial in Biotechnology and Cancer Research.

Table 3. Regression coefficients for the full regression model for Scopus narrow Biochemistry, Genetics and Molecular Biology fields (US only, first and last author gendered). Publication year regression coefficients are not shown. Fields are in descending order of female first author proportions in 2018.

Narrow field	Model 1					
	Female first	Female last	2 auth	3 auth	4 auth	5+ auth
Endocrinology	0.022***	0.009	0.081***	0.135***	0.173***	0.261***
Developmental Biology	-0.012*	-0.009	0.109***	0.132***	0.157***	0.261***
Aging	-0.006	0.025*	0.154***	0.235***	0.278***	0.353***
Cancer Research	-0.012**	-0.015**	0.327***	0.405***	0.432***	0.571***
Physiology	-0.004	-0.009	0.127***	0.141***	0.174***	0.254***
Genetics	-0.034***	-0.029***	0.168***	0.192***	0.207***	0.305***
Clinical Biochemistry	0.019**	0.005	0.211***	0.277***	0.323***	0.406***
Cell Biology	-0.017***	-0.025***	0.281***	0.325***	0.352***	0.447***
Molecular Biology	-0.014***	-0.017***	0.195***	0.231***	0.256***	0.333***
Molecular Medicine	0.003	-0.014*	0.304***	0.354***	0.387***	0.454***
Biochemistry	-0.013***	-0.015***	0.240***	0.274***	0.294***	0.364***
Structural Biology	-0.027**	-0.018	0.119***	0.147***	0.178***	0.243***
Biotechnology	-0.017*	-0.015	0.442***	0.487***	0.526***	0.630***
Biophysics	-0.021***	-0.028***	0.249***	0.277***	0.314***	0.354***

\* p<0.05; \*\* p<0.01; \*\*\* p<0.001

Table 4. Regression coefficients for three reduced regression models for Scopus narrow Biochemistry, Genetics and Molecular Biology fields (US only, first and last author gendered), ignoring team size. Publication year regression coefficients are not shown. Fields are in descending order of female first author proportions in 2018.

Narrow field	Model 2		Model 3	Model 4
	Female 1st	Female last	Female 1st	Female last
Endocrinology	0.030***	0.002	0.030***	0.006
Developmental Biology	-0.006	-0.015**	-0.008	-0.016**
Aging	0.004	0.013	0.006	0.014
Cancer Research	-0.001	-0.023***	-0.004	-0.023***
Physiology	0.000	-0.015**	-0.001	-0.015**
Genetics	-0.021***	-0.031***	-0.024***	-0.035***
Clinical Biochemistry	0.023***	-0.001	0.023***	0.002
Cell Biology	-0.009**	-0.029***	-0.011***	-0.030***
Molecular Biology	-0.007**	-0.019***	-0.009***	-0.020***
Molecular Medicine	0.009	-0.016*	0.008	-0.015*
Biochemistry	-0.003	-0.015***	-0.004	-0.015***
Structural Biology	-0.017	-0.016	-0.019	-0.019
Biotechnology	0.009	-0.023*	0.006	-0.022*
Biophysics	-0.004	-0.027***	-0.006	-0.028***

\* p<0.05; \*\* p<0.01; \*\*\* p<0.001

## Limitations

There are many limitations to this study in terms of factors that were impractical to consider. The gender differences found may be influenced by the Scopus classifications, especially for fields that fit poorly or articles in general journals. Internationally collaborative research is important and omitting this is a limitation, despite its necessity to avoid spurious second order gender effects. Citations do not reflect all aspects of research impact and it is possible that they better reflect the impact of males (Thelwall, 2018). Researcher seniority is ignored and so the results may be influenced by second order effects due to a greater share of males amongst senior academics. Other second order gender effects are also possible, such as males potentially publishing open access more (Zhu, 2017), perhaps conferring a citation advantage. Most importantly, any relationship between citation bias and field participation rates ignores many other factors known to differentially influence female participation rates between fields, such as support from existing faculty, alignment with career goals, societal expectations, and opportunities for relevant skills development in childhood (Ceci, Ginther, Kahn, & Williams, 2014).

## Discussion

The relatively low proportions of females in highly quantitative and computer-related subjects (Ceci, Ginther, Kahn, & Williams, 2014; Cheryan, Ziegler, Montoya, & Jiang, 2017) may account for their underrepresentation in three areas here: Biophysics, Biotechnology, and Structural Biology. Although Biotechnology was found to be more female globally in a previous study (Holman, Stuart-Fox, & Hauser, 2018), this seems to be a definitional issue because bioinformatics was categorised separately in that study and was male dominated. In support of this conclusion, biotechnology software code authors are known to be male dominated, especially for senior team members (Russell, Johnson, Ananthan, Harnke, & Carlson, 2018).

The small female majorities in Developmental Biology and Endocrinology suggests that they are relatively attractive STEM subjects for females. Psychology research suggests that females are more likely to choose a career with a goal of helping society (Diekmann, Steinberg, Brown, Belanger, & Clark, 2017), so these fields may be perceived in this way.

The gender disparities in team size (female first and male last authors tending to be in larger teams) are reasonably consistent. For last authors, this is presumably a second order effect of seniority, with senior males being more likely to be last authors. For first authors, senior authors may occasionally write solo articles, decreasing the team size average for male first authors (and last authors). Team size is a substantial factor for citation rates in some fields, associating with increases between 20% (Physiology) and 50% (Biotechnology) in the ratio of logged citations to the world average for the field(s) and year of publication. It is not clear whether this is a causal relationship due to larger teams creating better research or at least attracting more citations for their papers by generating a larger audience. Alternatively, large team research may coincidentally be in high citation micro-specialisms.

## Conclusions

The tendency for a small male citation advantage overall 1996-2014 for biochemistry, genetics and molecular biology journal articles contrasts with an overall 0.2% female US first author citation advantage for journal articles published in 2014 (Thelwall, 2018), mitigating

against citation advantages being able to account for the relatively high proportion of females in this category. There is also not a pattern for narrow fields within the 14 analysed here to have a higher proportion of females if they have a female citation advantage, or the opposite. For example, the two female fields in 2018, Developmental Biology and Endocrinology, have differencing gender effects signs in the main regression model, suggesting that citation-related gender effects are unlikely to have been a substantial factor in their recruitment and retention of females.

## References

- de Araújo, C. G. S., de Oliveira, B. R. R., de Oliveira Brito, L. V., da Matta, T. T., Viana, B. F., de Souza, C. P., & da Matta Mello Portugal, E. (2012). Two-year citations of JAPPL original articles: evidence of a relative age effect. *Journal of Applied Physiology*, 112(9), 1434-1436.
- Ceci, S. J., Ginther, D. K., Kahn, S., & Williams, W. M. (2014). Women in academic science: A changing landscape. *Psychological Science in the Public Interest*, 15(3), 75-141.
- Cheryan, S., Ziegler, S. A., Montoya, A. K., & Jiang, L. (2017). Why are some STEM fields more gender balanced than others? *Psychological Bulletin*, 143(1), 1-35.
- Diekmann, A. B., Steinberg, M., Brown, E. R., Belanger, A. L., & Clark, E. K. (2017). A goal congruity model of role entry, engagement, and exit: Understanding communal goal processes in STEM gender gaps. *Personality and Social Psychology Review*, 21(2), 142-175.
- Elsevier (2017). Gender in the global research landscape. Retrieved from: [https://www.elsevier.com/data/assets/pdf\\_file/0008/265661/ElsevierGenderReport\\_final\\_for-web.pdf](https://www.elsevier.com/data/assets/pdf_file/0008/265661/ElsevierGenderReport_final_for-web.pdf)
- Garg, K. C., & Kumar, S. (2014). Scientometric profile of Indian scientific output in life sciences with a focus on the contributions of women scientists. *Scientometrics*, 98(3), 1771-1783.
- Gazni, A., & Didegah, F. (2011). Investigating different types of research collaboration and citation impact: a case study of Harvard University's publications. *Scientometrics*, 87(2), 251-265.
- Guglielmi, G. (2019). Eastern European universities score highly in university gender ranking. *Nature News*, 10.1038/d41586-019-01642-4
- Holman, L., Stuart-Fox, D., & Hauser, C. E. (2018). The gender gap in science: How long until women are equally represented? *PLoS Biology*, 16(4), e2004956.
- Hopkins, A. L., Jawitz, J. W., McCarty, C., Goldman, A., & Basu, N. B. (2013). Disparities in publication patterns by gender, race and ethnicity based on a survey of a random sample of authors. *Scientometrics*, 96(2), 515-534.
- Larivière, V., Desrochers, N., Macaluso, B., Mongeon, P., Paul-Hus, A., & Sugimoto, C. R. (2016). Contributorship and division of labor in knowledge production. *Social Studies of Science*, 46(3), 417-435.
- Larivière, V., Gingras, Y., Sugimoto, C. R., & Tsou, A. (2015). Team size matters: Collaboration and scientific impact since 1900. *Journal of the Association for Information Science and Technology*, 66(7), 1323-1332.
- Larivière, V., Ni, C., Gingras, Y., Cronin, B., & Sugimoto, C. R. (2013). Bibliometrics: Global gender disparities in science. *Nature News*, 504(7479), 211-213.

- Levitt, J. & Thelwall, M. (2008). Is multidisciplinary research more highly cited? A macro-level study. *Journal of the American Society for Information Science and Technology*, 59(12), 1973-1984.
- NCES (2019). Table 315.80. Full-time and part-time faculty and instructional staff in degree-granting postsecondary institutions, by race/ethnicity, sex, and program area: Fall 1998 and fall 2003, [https://nces.ed.gov/programs/digest/d17/tables/dt17\\_315.80.asp](https://nces.ed.gov/programs/digest/d17/tables/dt17_315.80.asp)
- O'Dea, R. E., Lagisz, M., Jennions, M. D., & Nakagawa, S. (2018). Gender differences in individual variation in academic grades fail to fit expected patterns for STEM. *Nature Communications*, 9(1), 3777.
- Russell, P. H., Johnson, R. L., Ananthan, S., Harnke, B., & Carlson, N. E. (2018). A large-scale analysis of bioinformatics code on GitHub. *PLoS ONE*, 13(10), e0205898.
- Sax, L. J., Lim, G., Lehman, K., & Lonje-Paulson, L. (2018). Reversal of the Gender Gap: The biological sciences as a unique case within science, technology, engineering, and mathematics (STEM). *Journal of Women and Minorities in Science and Engineering*, 24(4), 291-324.
- Thelwall, M., Bailey, C., Tobin, C. & Bradshaw, N. (2019). Gender differences in research areas, methods and topics: Can people and thing orientations explain the results? *Journal of Informetrics*, 13(1), 149-169.
- Thelwall, M. (2017). Three practical field normalised alternative indicator formulae for research evaluation. *Journal of Informetrics*, 11(1), 128–151.
- Thelwall, M. (2018). Do females create higher impact research? Scopus citations and Mendeley readers for articles from five countries. *Journal of Informetrics*, 12(4), 1031-1041.
- Tscharntke, T., Hochberg, M. E., Rand, T. A., Resh, V. H., & Krauss, J. (2007). Author sequence and credit for contributions in multiauthored publications. *PLoS Biology*, 5(1), e18.
- Waltman, L., van Eck, N. J., van Leeuwen, T. N., Visser, M. S., & van Raan, A. F. (2011). Towards a new crown indicator: An empirical analysis. *Scientometrics*, 87(3), 467-481.
- Yohai, V. J. (1987). High breakdown-point and high efficiency robust estimates for regression. *The Annals of Statistics*, 15(2), 642-656.
- Zhu, Y. (2017). Who support open access publishing? Gender, discipline, seniority and other factors associated with academics' OA practice. *Scientometrics*, 111(2), 557-579.
- Zitt, M. (2012). The journal impact factor: Angel, devil, or scapegoat? A comment on JK Vanclay's article 2011. *Scientometrics*, 92(2), 485-503.