Check for updates

Citation inequality and the Journal Impact Factor: median, mean, (does it) matter?

Tobias Kiesslich, et al. *[full author details at the end of the article]*

© The Author(s) 2020

Abstract

Skewed citation distribution is a major limitation of the Journal Impact Factor (JIF) representing an outlier-sensitive mean citation value per journal The present study focuses primarily on this phenomenon in the medical literature by investigating a total of $n = 982$ journals from two medical categories of the Journal Citation Report (JCR). In addition, the three highest-ranking journals from each JCR category were included in order to extend the analyses to non-medical journals. For the journals in these cohorts, the citation data (2018) of articles published in 2016 and 2017 classified as citable items (CI) were analysed using various descriptive approaches including e.g. the skewness, the Gini coefficient, and, the percentage of CI contributing 50% or 90% of the journal's citations. All of these measures clearly indicated an unequal, skewed distribution with highly-cited articles as outliers. The %CI contributing 50% or 90% of the journal's citations was in agreement with previously published studies with median values of 13–18% CI or 44–60% CI generating 50 or 90% of the journal's citations, respectively. Replacing the mean citation values (corresponding to the JIF) with the median to represent the central tendency of the citation distributions resulted in markedly lower numerical values ranging from −30 to −50%. Up to 39% of journals showed a median citation number of zero in one medical journal category. For the two medical cohorts, median-based journal ranking was similar to mean-based ranking although the number of possible rank positions was reduced to 13. Correlation of mean citations with the measures of citation inequality indicated that the unequal distribution of citations per journal is more prominent and, thus, relevant for journals with lower citation rates. By using various indicators in parallel and the hitherto probably largest journal sample, the present study provides comprehensive up-to-date results on the prevalence, extent and consequences of citation inequality across medical and all-category journals listed in the JCR.

Keywords Journal Impact Factor · Citation distribution · Skewed distribution · Journal quality · Article citedness

Introduction

The Journal impact factor (JIF) has been introduced by Garfield (1955) partly based on previous work by Gross and Gross (1927) and initially intended to assist librarians for deciding which journals to purchase for their institution (Garfield 2006). While its calculation is simply the arithmetic mean (Pang 2019) by division of the number of citations a journal receives in a given year (numerator) by the number of papers which received these citations and are published by that journal in the two preceding years (denominator), the JIF is probably the most controversial metric. The main reason for that is due to its widespread use as a proxy of research quality or research performance in the assessment of individual authors, departments or academic institutions (Adler et al. 2009; Seglen 1989), see also (Opthof 1997; Opthof and Wilde 2009) for a further discussion on the use of citation data to evaluate research. As pointed out by several colleagues including concerns of his inventor ["used inappropriately as surrogates in evaluation exercises", (Garfield 1996)], such usage is mostly inappropriate [e.g. Simons (2008), McKiernan et al. (2019), Casadevall and Fang (2014)] as the JIF of a journal does not predict the citedness of individual articles published in the respective journal [e.g. Adler et al. (2009), Opthof (1997), Opthof et al. (2004)]. Several initiatives have been instigated discussing the potentially harmful effects of such practices and providing recommendations for sensible and responsible use of metrics such as the JIF—most prominently the 'San Francisco Declaration on Research Assessment' (DORA),[1] the 'Leiden Manifesto' (Hicks et al. 2015), and 'the Metric tide' report (Wilsdon et al. 2015). Besides several other general lines of critique [see e.g. Seglen (1998), Larivière et al. 2016, Larivière and Sugimoto 2019; Casadevall and Fang 2014; Glänzel and Moed 2002)], the lack of correlation between JIF and article citedness is primarily based on the (highly) asymmetric distribution of citations to a journal's published articles. In particular, this means that (1) the citations received by a journal are not equally distributed among the papers this journal has published, and that (2) the majority of papers in a given journal are cited infrequently compared to a few highly-cited 'outliers'. About 30 years ago, Seglen (1989, 1992) has provided convincing evidence for this "skewness of science" in terms of citation distributions—recently confirmed by Zhang et al. (2017). Altogether, the resulting arithmetic mean is strongly influenced by a minority of highly-cited papers and does not adequately reflect the "average" citation rate a particular journal is characterised by Campbell (2008). Consequently, the median value of citations to a journal was suggested to be more appropriate to express the journal's citedness [e.g. Editor(s) (2011), Opthof (2019), Pulverer (2013, 2015), Weale et al. (2004)].

As summarised in the discussion section, several previous studies investigated the distribution of citations using a variety of approaches for such descriptive analyses and rather heterogeneous selections of investigated journals or group of journals. The current study aimed to provide a comprehensive up-to-date analysis of the citation distribution characteristics (citation inequality, skewness) based on three independent cohorts of journals listed in the recent Journal Citation Report (JCR 2018): journals of two complete medical categories, i.e. 'Medicine, Research & Experimental' and 'Medicine, General & Internal' and the three best-ranking journals in each JCR category (further referred to as 'Med-R&E', 'Med-G&I', and 'Top 3', respectively). The two medical categories were chosen due to

---

Table 1 Journal cohort characteristics	Cohort (JCR category)	Med-R&E	Med-G&I	Top 3
	JCR Edition (2018)	SCIE	SCIE	SCIE & SSCI
	Journals analysed (n)	135	159	688
	JIF			
	Range[a]			
	Min	0.181	0.071	0.828
	Max	30.641	70.670	223.679
	Numerator (n citations)[a]			
	Min	15	8	18
	Max	9376	46,289	73,952
	Denominator (n, CI)[a]			
	Min	27	27	3
	Max	5073	6874	5070
	Total CI analysed (n)	56,287	52,428	258,314

[a]Minimum and maximum values for the JIF, its numerator and denominator represent the range of these characteristics for individual journals in each of the three cohorts

CI citable items, *JCR Journal Citation Report*, *JIF* Journal Impact Factor, *Med-G&I* Medicine, General & Internal, *Med-R&E* Medicine, Research & Experimental, *SCIE* Science Citation Index Expanded, *SSCI* Social Science Citation Index

the perceived high relevance of the JIF and the discussion of its use and misuse especially in (bio-)medical sciences. The Top 3 cohort was included to provide an extension of the results to non-medical journals by analysis of the "best" three journals in all JCR categories. The claim of novelty and comprehensiveness of this study is based on the combined use of various previously reported approaches to describe and quantify the skewness of the citation distribution in a large dataset comprising a total of 982 journals. Besides the aim of investigating the prevalence of citation skewness in the two complete general medical journal categories, the third category was included to study whether the phenomenon of skewed citation distributions for individual journals or (small) cohorts of journals within a JCR category could be extended to all journals and subject categories currently indexed in the JCR and thus having been assigned a JIF.

Data and methods

Journal cohorts

Journal datasets analysed in the current study comprised two complete medical categories in the SCIE Edition of the 2018 *Journal Citation Report* (JCR)—'Medicine, Research & Experimental' and 'Medicine, General & Internal'—as well as the three highest ranking journals (JIF-based ranking) from all categories of both 2018 JCR editions, i.e. the Science Citation Index Expanded (SCIE) and the Social Science Citation Index (SSCI). Table 1 lists the basic characteristics of these cohorts.

All journals in the respective cohort were included for further analysis except one journal in each of the two medical categories for which no JIF has been published in the 2018 JCR. The third cohort (Top-3) comprises journals from all SCIE and SSCI categories ($n = 236$ categories). In case individual categories are listed under the same name in both editions of the JCR (SCIE and SSCI), the list of journals was treated as if it was one category followed by selection of the top 3 journals. This was the case for the following seven categories: 'Green & Sustainable Science & Technology', 'History & Philosophy of Science', 'Nursing', 'Psychiatry', 'Public, Environmental & Occupational Health', 'Rehabilitation', 'Substance Abuse'. Therefore, a total of $n = 229$ categories (236 minus 7) were included in this Top 3 cohort. In the category 'Nursing', four journals were included because the third rank was occupied by two journals with identical JIF values. Consequently, $n = 688$ journals were included in this cohort (Top-3): 229 categories × 3 journals = 687; plus one journal from the 'Nursing' category = 688. To ensure consistency, the three highest-ranking journals (based on JIF) in the two medical categories were also included in the Top 3 cohort. Therefore, this cohort includes all journals that occupy the first three rank positions in the respective JCR categories.

Data collection

Using the web interface of the Clarivate's JCR (https://jcr.clarivate.com) via the authors' institution's subscription, the citation data for each journal were retrieved using the following procedure: in the 'Browse by Journal' option of the JCR, the list of journals was restricted to each particular category by 'Select Categories'. Afterwards, each journal (either the complete list for the two medical categories or the top 3 journals based on JIF ranking) was opened in a new tab of the web browser providing each journal's 'Journal Profile'. The complete list of articles (citable items (CI)) was exported to a *.csv file as outlined in detail in Online Resource 1 including representative screenshots. Beside basic bibliographic information on each article (citable items in 2016 and 2017), these files contained the number of citations the CI have received in 2018. The data were retrieved in January 2020 covering the data of the 2018 JCR and reflect the content of the database at that point in time. For further analysis, the journal names and the citations for each CI per journal were compiled into a single Excel file for each of the three cohorts. The published JIF for the journals in the three cohorts was retrieved by downloading the latest "JCRs data" ("JCR SSCI 2018 Metrics" and "JCR SCI 2018 Metrics" published on Nov 8, 2019).

Descriptive analysis

The distribution of citations for each journal was assessed using several approaches: (i) Lorenz curves showing the cumulative percentage of citations a journal received in 2018 versus the cumulative percentage of articles (CI) in 2016 and 2017 these citations refer to, (ii) the percentage of CI with $n = 0$ citations, (iii) the percentage of CI achieving 50 or 90% of the journal's total citations (50/90% cumulative citations threshold), (iv) the percentage of citations generated by the 50% most-cited CI per journal, (v) the percentage of CI with citations greater than the corresponding mean citation rate, (vi) the skewness of the distribution, and, (vii) the Gini coefficient. In brief, for (i), the raw citation data (sorted by citations in descending order) and the number of CI per journal were transformed to relative percentages using Microsoft Office Excel (v. 2016; Redmond, WA, USA) for the double-cumulative plots (Lorenz curves). The procedure is described in detail in Online Resource

2. For (ii)–(iii) and (v), the Excel function COUNTIF was used to determine the number of CI fulfilling the respective criterion. For (iv), the Excel function SUMIF was used in combination with the OFFSET function—the latter specifying the range of the most-cited 50% citable items. The skewness of the citation distribution (vi) was calculated based on the raw citation data per journal using the STATISTICS ON COLUMNS function of OriginPro 2020b (OriginLab Corp., Northampton, MA, USA) by the formula given in Eq. 1

$$skewness = \frac{n}{(n-1)(n-2)} \sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{sd} \right)^3 \tag{1}$$

with n is the number of values ($x_1, x_2,..., x_n$), and sd is the standard deviation (OriginLab manual, https://www.originlab.com/doc/X-Function/ref/moments). For (vii), the Gini coefficient as a measure of inequality (De Maio 2007; Gini 2005) was calculated based on the double cumulative percentages (% citations versus %CI): for each journal's Lorenz curve, the area (A_{ue}) was calculated using the OriginPro INTEGRATE function (mathematical areas, i.e. the algebraic sum of trapezoids). The Gini coefficient (G) was subsequently calculated as shown in Eq. 2

$$G = \frac{A_{ue} - A_e}{A_e} \tag{2}$$

with A_{ue} is the area of the putative unequal distribution (i.e. journal's citation distribution) and A_e is the area below the theoretical line of equality (45° line corresponding to an area of 5000 in the double cumulative plot (0–100% on each axis)). Division by A_e normalises the results yielding Gini coefficients in the range between 0 (completely equal distribution) and 1 (highest possible inequality). For a numerical example of this procedure, see Online Resource 2. Similar to the median citation rate, the mean citation rate was calculated manually from the raw data. As further outlined in the results section, the "mean citations" correspond to the published JIF but including only citations to "citable items" listed in the denominator.

Statistics and data visualisation

Basic calculations were performed in Microsoft Office Excel (v. 2016) and data were visualised using OriginPro 2020b and Corel Designer 2018 (Corel Corp., Ottawa, Canada). Box plots show data points (left) and median values (horizontal line), the 25–75% percentile (box, right) as well as the min–max values. Spearman correlation analysis was used to investigate the relationship between the mean citations and various measures of citation inequality considering p values < 0.05 (< 0.01) as (highly) statistically significant.

Results

Citation distributions

Figure 1a–c show the double cumulative plots (% citations versus %CI) for each journal in the categories Med-R&E (Fig. 1a), Med-G&I (Fig. 1b) and the Top 3 journals from all JCR categories (Fig. 1c). This demonstrates the skewness of the citation distribution since the Lorenz curves of all journals in the two medical categories significantly deviate from
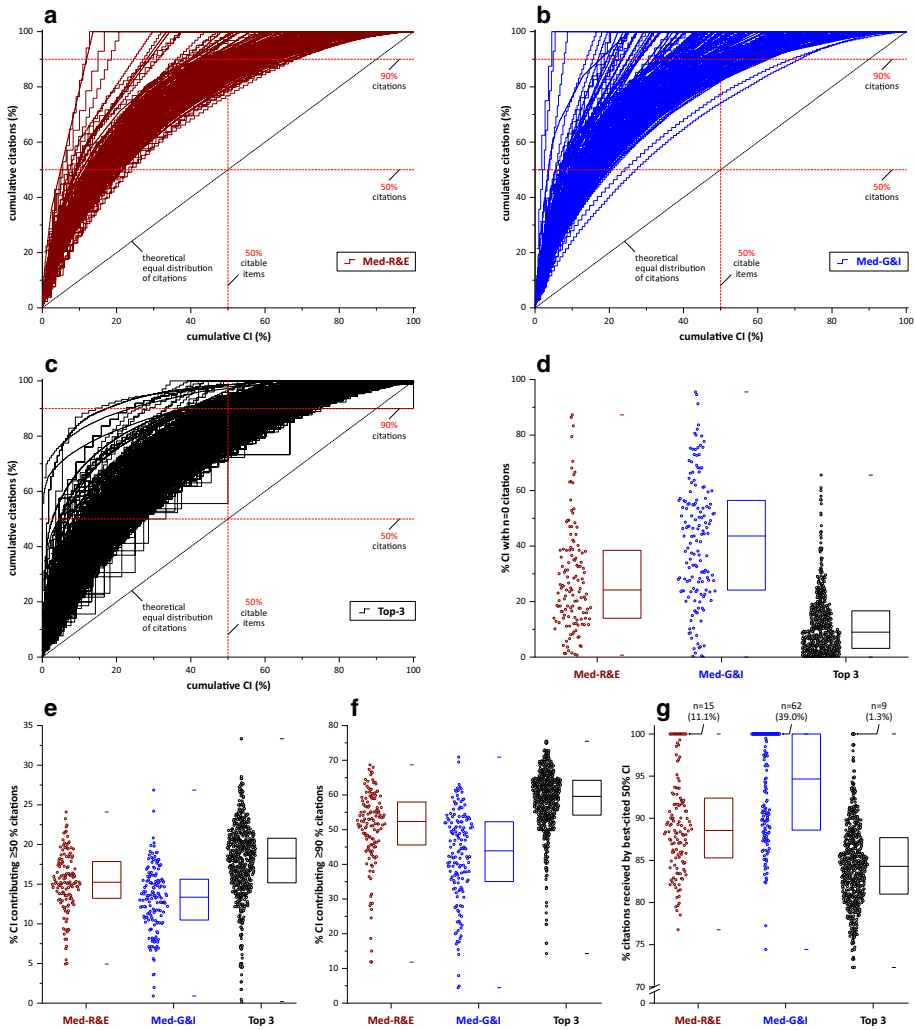
Fig. 1 Citation distributions. **a–c** Lorenz curves showing the cumulative % of citations versus the cumulative % of CI for the three journal cohorts Med-R&E, Med-G&I and Top 3, respectively. Horizontal dashed red lines indicate the 50 or 90% cumulative citations threshold. The vertical dashed line indicates the 50% best-cited CI threshold. **d** %CI with $n=0$ citations. **e/f** %CI contributing $\geq 50\%/\geq 90\%$ citations to the journal. **g** % citations generated by 50% of the best-cited papers. The box plots in **d–g** show the median (horizontal line), the 25–75% interquartile range, and maximum/minimum values. *Abbreviations*: CI = citable items, Med-G&I = Medicine, General & Internal, Med-R&E = Medicine, Research & Experimental, Top 3 = the three highest-ranking journals (JIF-based ranking) from all *Journal Citation Reports* categories

the theoretical 45° line of equality. Figure 1d shows the median percentage of CI receiving $n=0$ citation (thus not contributing to the JIF) which is 24.2% and 43.6% of papers for the medical cohorts Med-R&E and Med-G&I, respectively. Noticeably, in these JCR categories, several journals have published up to 87% (Med-R&E) and 96% (Med-G&I) citable items that have never been cited within the JIF window. For the Top 3 cohort, these values are considerably lower, i.e. a median of 9.0% uncited items and the journal with highest

proportion of uncited publications includes 66% uncited items. In this cohort, i.e. the top 3 journals from all JCR categories based on JIF ranking, the majority of journals have $< 20\%$ CI with $n = 0$ citations.

The inequality of the citation distributions is further illustrated by the %CI published in a journal required for generation of $\geq 50\%$ (Fig. 1e) or $\geq 90\%$ of its total citations (Fig. 1f; in Fig. 1a–c, these thresholds are indicated by dashed horizontal lines). For the medical JRC categories Med-R&E and Med-G&I, a median of 15.2% (52.4%) and 13.4% (43.9%) citable items are needed for generation of $\geq 50\%$ ($\geq 90\%$) of the total citations to the journal, respectively. In case of the Top 3 cohort, a median of 18.3% (59.6%) of citable items generate $\geq 50\%$ ($\geq 90\%$) of the journal's total citations. For all three cohorts, several journals show a one-digit % range of citable items which contribute $\geq 50\%$ of all citations. Maximum values are about 30% for CI contributing $\geq 50\%$ of citations and about 70% for CI contributing $\geq 90\%$ of citations in the three journal cohorts. The majority of journals groups around the median value of %CI required for $\geq 50\%$ or $\geq 90\%$ of citations (Fig. 1e, f). A different approach calculates the % of total citations generated by the most cited 50% of citable items (Fig. 1g): here, a median of 88.6/94.7/84.3 of all citations were received by the most cited 50% of papers for the three categories, respectively. In the cohorts Med-R&E, Med-G&I and Top-3, fifteen (11.1%), 62 (39.0%) and nine (1.3%) journals even generate all citations (100%) with their 50% most-cited articles, respectively.

Measures of inequality

The numerator of the published ('official') JIF might differ from the actual sum of citations generated only by the CI in the JIF window. This discrepancy can be explained by unmatched citations (Larivière et al. 2016) and the known asymmetry between the JIF's numerator and denominator [e.g. (Glänzel and Moed 2002)], i.e. a journal can acquire citations to ('non-citable') content that is, however, not counted (classified as CI) in the denominator. As shown in Online Resource 3, this is true for almost all journals in the cohorts investigated in this study: except for six and three journals in the categories Med-R&E and Med-G&I, respectively, where the difference between the JIF numerator and the sum of citations to CI is zero. For all other journals, the JIF numerator counts more citations than those received by publications classified as citable items. The median % reduction when using the actual sum of citations instead of the JIF numerator is $-5.3\%/-10.6\%/-5.9\%$ for the three cohorts, i.e. Med-R&E, Med-G&I and Top 3, respectively. While most journals show a % reduction in that order of magnitude, extreme examples in the three cohorts are journals having 28.8%/47.8%/80.5% fewer citations and, consequently, also fewer mean citations than indicated by JIF numerator and the published JIF, respectively (see Online resource 3d).

Therefore, for subsequent analyses, not the published JIF from the JCR database was used, instead this indicator was calculated manually as the arithmetic mean of the raw citation data for each citable item (CI) per journal. Thus, only citations unambiguously matched to citable items were used for calculating mean citation values in the current study in line with what had been previously argued (Larivière et al. 2016). This approach was chosen since the median value (for citations per journal) is not included in the downloadable version of the JCR: therefore, in order to ensure consistency and comparability between the measures of central tendency, both mean and median citations were calculated manually for all subsequent analyses. Consequently, the unequal distribution of citations to the CI per journal can further be characterised by how many
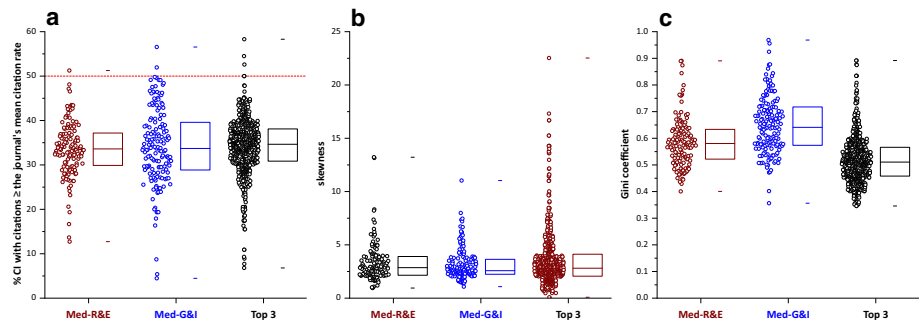
Fig. 2 Measures of inequality of citation distributions. **a** % CI with citations equal or greater than the JIF. **b**, **c** Skewness and Gini coefficients of the citations per journal, respectively. The box plots show the median (horizontal line), the 25–75% interquartile range, and maximum/minimum values. *Abbreviations*: CI = citable items, Med-G&I = Medicine, General & Internal, Med-R&E = Medicine, Research & Experimental, Top 3 = the three highest ranking journals (JIF-based ranking) from all *Journal Citation Reports* categories

citable items (%CI) per journal actually receive citations that are equal to or higher than the journal's arithmetic citation mean (corresponding to the JIF, but without citations to non-citable items): as shown in Fig. 2a, only a median of 33.6%/33.7%/34.7% citable items receive at least as many citations as the mean number of citations in the three cohorts, respectively.

Notably, only one journal in Med-R&E, two journals in Med-G&I and five journals in the Top 3 cohort comprise equal or more than 50% of CI that receive citations equal or more than the mean citation for the respective journal—in case of a symmetric, equal distribution of citations, half of each journal's CI (50%) should receive citations equal or higher than the journal's JIF. It seems worth mentioning that using the published ('official') JIF for this calculation (i.e., %CI with citations ≥ JIF) would result in even smaller numbers of journals.

The asymmetry of a distribution can be additionally quantified using the skewness: in case the right tail is longer (right-skewed, left-leaning distribution), a positive skew is expected (Henderson 2006). For the current data set, this is true for all journals as shown in Fig. 2b: the min–max skewness values are 0.9–13.2/1.1–11.0/0.1–22.5 for the Med-R&E, Med-G&I and Top 3 cohorts, respectively. The median skewness for the three groups of journals is rather similar with 2.9/2.6/2.8, respectively.

Finally, the Gini coefficient as a widely used measure of inequality in economic studies has been calculated for the journals' citations in the three cohorts. This measure is defined as the difference between the area under the curve of the theoretical equal distribution (45° straight line in the Lorenz curve plots; compare Fig. 1a-c and Online Resource 2) and the area below the actual distribution of the variable of interest (e.g. income across the population). While a Gini coefficient of zero indicates complete equality, a value of 1 indicates an entirely unequal distribution of a value (De Maio 2007)—that would be a single CI receiving all citations to a journal. For the three cohorts in the current study (Fig. 2c), the Gini coefficients' medians are 0.58/0.64/0.51 ranging between min–max values of 0.40–0.89/0.36–0.97/0.35–0.89, respectively. In all cohorts, extreme examples of journals have Gini coefficients of up to 0.9 or higher (Fig. 2c).

The relation of the mean citation values and the above-mentioned measures of unequal citation distribution is graphically shown in Online Resource 4 and summarised in Table 2.

Table 2 Correlation of mean citations versus measures of citations' inequality

Cohort	% citable items				Skewness	Gini coefficient
	Contributing		With 0 citations	With citations ≥ mean citations		
	≥ 50% citations	≥ 90% citations				
Med-R&E	**0.6092**	**0.8259**	**−0.9581**	0.0653	0.1070	**−0.7624**
	< 0.0001**	**< 0.0001****	**< 0.0001****	0.4517	0.2169	**< 0.0001****
Med-G&I	**0.7865**	**0.9352**	**−0.9830**	0.1069	0.1217	**−0.8975**
	< 0.0001**	**< 0.0001****	**< 0.0001****	0.1797	0.1265	**< 0.0001****
Top 3	**0.2563**	**0.5526**	**−0.8607**	**−0.2190**	**0.1778**	**−0.4091**
	< 0.0001**	**< 0.0001****	**< 0.0001****	**< 0.0001****	**< 0.0001****	**< 0.0001****

Shown are the Spearman correlation coefficients and the corresponding p values

JIF Journal Impact Factor, *Med-G&I* Medicine, General & Internal, *Med-R&E* Medicine, Research & Experimental

**p < 0.01 is considered highly statistically significant (bold)

While there is a heterogeneously strong and significant positive correlation between mean citations and the %CI generating ≥ 50 or ≥ 90% of total citations in all three cohorts, its association with %CI with $n = 0$ citations is negative and highly significant in all cohorts. The relationship between mean citations and the percentage of CI reaching equal or more citations than the journal's mean citation is—although significant in the Top 3 cohort—rather weak (see also Online Resource 4d). The same interpretation applies to the association of mean citations and the skewness of the citation distribution. A moderate and significant correlation can be observed for the mean citations and the Gini coefficient: here, the journals' mean citation values are inversely proportional to the Gini coefficient.

Taken together, several mathematical approaches confirm a general, unequal and skewed distribution of citations for all journals analysed.

Mean versus median

Given the asymmetric distribution of the citations per journal, the median seems to be more appropriate to represent the central tendency. Therefore, median citations were calculated for each journal and compared to mean-based citation rates for the journals of the three cohorts in this study (Fig. 3a-c): for the vast majority of journals, the median is considerably lower than the mean value—with very few exceptions where the opposite is the case ($n = 1, 1, 4$ journals for the Med-R&E, Med-G&I and Top 3 cohorts, respectively). For some journals ($n = 4, 1, 15$ for the Med-R&E, Med-G&I and Top 3 cohorts, respectively) calculation of the median citation gives a non-integer value (*.5): these are cases of an even number of CI where the arithmetic mean of the two central values is non-integer.

As summarised in Fig. 3d, the numerical values of the mean citations are reduced by a median of minus 36.3/50.1/31.4% when median values are used instead for characterising the three cohorts, respectively. Furthermore, the median citation rate drops to zero for $n = 15/62/9$ journals (11.1/39.0/1.3% of journals) in the three cohorts, respectively. By definition of the median, these values correspond to the percentages of journals generating all of their citations with 50% (or less) of their most-cited papers as
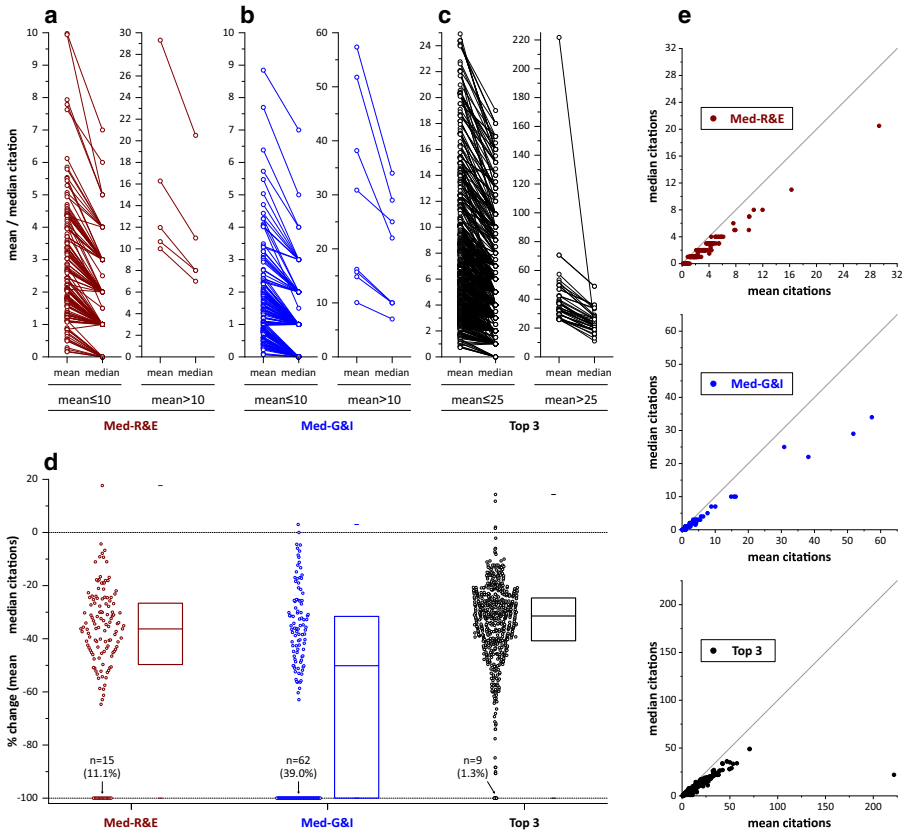
Fig. 3 Mean versus Median citations (I). **a–c** Mean (left) versus median citations (right) is shown for the three journal cohorts Med-R&E, Med-G&I and Top 3, respectively. For better readability, the data are split into two diagrams using a JIF threshold of 10 or 25 for the two medical and the Top 3 cohort of journals, respectively. **d** % reduction of the numerical value if median instead of mean citations are used for each journal. **e** Correlation of the median and mean citation values per journal. The box plots show the median (horizontal line), the 25–75% interquartile range, and maximum/minimum values. *Abbreviations*: Med-G&I=Medicine, General & Internal, Med-R&E=Medicine, Research & Experimental, Top 3=the three highest ranking journals (JIF-based ranking) from all *Journal Citation Reports* categories

shown in Fig. 1g. The correlation between mean and median citations shown in Fig. 3e furthermore demonstrate an increasing difference in absolute numbers between these variables for journals with higher mean citation values (i.e. 'top journals' in terms of mean citation rates).

The median of citations each journal receives in the JIF window was further used to rank the journals in the two medical JRC categories: Fig. 4a (Med-R&E) and Fig. 4b (Med-G&I) compares the ranking of journals if their position was determined by the mean citations (left) or median citations (right). Using the median as the ranking criterion would reduce the number of rank positions to 13 with 1–3 instances with a non-integer (*.5) median citation increment. Especially ranking positions with a median citation rate of 0–4 are occupied by several journals.
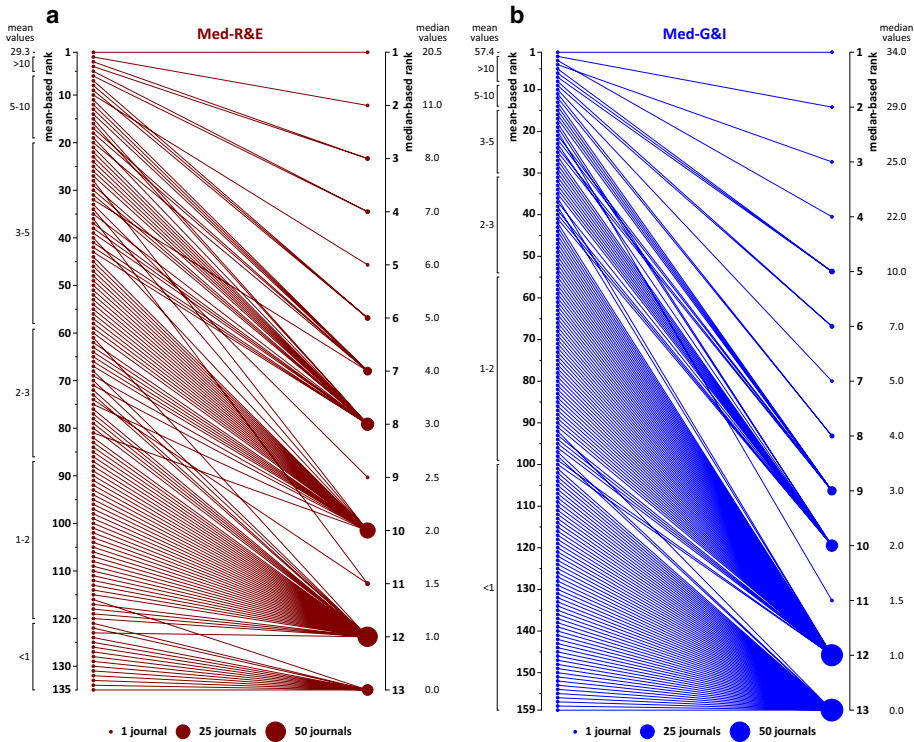
Fig. 4 Mean versus Median citations (II). **a**, **b** mean-based versus median-based ranking of journals in the medical categories Med-R&E and Med-G&I, respectively. The number of journals in the median-based ranking is coded by the area of the circle. *Abbreviations*: Med-G&I=Medicine, General & Internal, Med-R&E=Medicine, Research & Experimental

Discussion

This study investigated the citation distribution and its quantitative characteristics of three cohorts of journals: two complete JCR categories including journals with a general medical focus ('Research & Experimental', and 'General & Internal') as well as the three top-ranked journals from all categories of the JCR as a comparison. The analyses presented in this paper confirm a highly skewed distribution of citations per journal for all investigated cohorts—quantified by several measures to approach this phenomenon from different angles.

Table 3 provides an overview of previous studies reporting quantitative characteristics of citation distributions on a journal level [for citation skew on the article level and basically corresponding results, see e.g. (Albarrán et al. 2011; Bornmann and Leydesdorff 2017)]. For the first indicator, i.e. the percentage of citable items receiving less citations than indicated by the JIF, the present study obtained about 65% of CI for all three journal cohorts (corresponding to about 35% CI receiving more citation than the mean)—similar to previously reported values (Asaad et al. 2019; Larivière et al. 2016; Larivière and Sugimoto 2019). A symmetric distribution would result in about 50% CI below/above the mean (JIF) thus demonstrating the JIF to overestimate the real 'average' citation rate in virtually

Table 3 Literature review—summary of studies reporting quantitative characteristics of citation distributions on a journal level

Journal[a]/cohort	Journals	CI year	% citable items			Contributing		% citations received by 50% CI	Skewness	Gini coefficient	Reference
			With citations		With n=0 citations	50% citations	90% citations				
			<JIF	≥JIF							
Nat Mater	1	2008/2009				16%					(Editor(s) 2011)
Cardiovasc Res	1	1992–2000			34.3%	14.0±4.8%		90.3±6.1%			(Opthof et al. 2004)
Cardiovasc Res Circ Res	2	1992/1993				12%[b] 16%[b]		92%[b] 85%[b]			(Opthof and Coronel 2002)
Sao Paulo Med J Clinics Rev Assoc Med Bras	3	2007/2008			62%[b] 35%[b] 71%[b]						(Lustosa et al. 2012)
J Biol Chem, Biochem J, Biochim Biophys Acta	3	1983/1984				15%[c]		~90%[c]			(Seglen 1992)
J Biol Chem Int J Cancer Scand J Immunol	3	2000/2005 2000/2005 2000/2005				19%/18% 22%/21% 23%/24%		82%/86% 80%/79% 78%/77%		0.395[d] 0.532[d]	(Zhang et al. 2017)
Clin Pharmacol Ther Brit J Clin Pharmaco EUR J Clin Pharmacol	3[e]	1981/1991/1995/1996							+1.47 +2.16 +1.37		(Rostami-Hodjegan and Tucker 2001)

Table 3 (continued)

Journal[a]/cohort	Journals	CI year	% citable items		With n=0 citations	Contributing		% citations received by 50% CI	Skewness	Gini coefficient	Reference
			With citations			50% citations	90% citations				
			<JIF	≥JIF							
2016 JCR journals	n.d.	2014–2015		~30%							(Larivière and Sugimoto 2019)
3 High JIF Journals	9	2005–2007				16.1/12.0%[f]					(Falagas et al. 2010)
3 Moderate JIF Journals						21.7/17.6%[f]					
3 Low JIF Journals						15.9/14.6%[f]					
Vet Anaesth Analg	1	2007–2017				15–20%	~50%				(Pang 2019)
n=10 top-ranking	10	2014/2015				<13%	<40%				
n=10 low-ranking	10	2014/2015				15–20%	<60%				
Selection[g]	11	2013/2014	65–75%			15–25%					(Larivière et al. 2016)
30 Most cited cardiovascular journals	30	2015/2016		27.8% ±1.1%	9.0% ±1.1%						(Opthof 2019)
Cohort: Plastic Surgery	33	2016/2017	66%	34%	40%	12.6%					(Asaad et al. 2019)
Top 3 Journals[h]	3					15.2–20.2%					

Table 3 (continued)

Journal[a]/cohort	Journals	CI year	% citable items With citations <JIF	% citable items With citations ≥JIF	With n=0 citations	Contributing 50% citations	Contributing 90% citations	% citations received by 50% CI	Skewness	Gini coefficient	Reference
Cohort: Ortho- paedics	74	2013/2014	85%		64.5%						(Bozzo et al. 2017)
Immunology[i]	105	2000–2001			17.6% [10–32]	18% [15–21]					(Weale et al. 2004)
Surgery[i]	120				32.8% [21–50]	15% [13–18]					
Cohort: Med-R&E[i]	135	2016/2017	66.4% [62.8–70.1]	33.6% [29.9–37.2]	24.2% [14.0–38.5]	15.2% [13.2–17.9]	52.4% [45.6–58.0]	88.6% [85.3–92.4]	+2.9 [2.1–3.9]	0.58 [0.52–0.63]	this study
Cohort: Med-G&I[i]	159	2016/2017	66.3% [60.4–71.1]	33.7% [28.9–39.6]	43.6% [24.1–56.5]	13.4% [10.5–15.6]	43.9% [35.0–52.3]	94.7% [88.6–100.0]	+2.6 [2.2–3.6]	0.64 [0.57–0.72]	
Cohort: Top 3[i]	688	2016/2017	65.3% [61.9–69.1]	34.7% [30.8–38.1]	9.0% [3.1–16.6]	18.3% [15.1–20.7]	59.6% [54.1–64.2]	84.3% [81.0–87.7]	+2.8 [2.1–4.1]	0.51 [0.45–0.56]	

(a)…in case of individual journals, the journal title is given by the JCR abbreviation; (b)…estimated from the diagrams in (Opthof and Coronel 2002) and (Lustosa et al. 2012); (c)…mean values for the three journals; (d)…for 2005; (e)…n=30 randomly chosen papers per journal per year. (f)…for original articles/reviews, respectively; (g)…the 11 journals selected are both, multidisciplinary and subject-specific, have a JIF in the range between 3 to >30, and are published by seven different publishers (Larivière et al. 2016); (h)…the top 3 journals in the cohort of plastic surgery journals (Asaad et al. 2019); (i)…numbers in square brackets indicate the 25–75% interquartile range (IQR)

CI citable items, Med-G&I Medicine, General & Internal, Med-R&E Medicine, Research & Experimental, n.d. not disclosed, Top 3 the three highest ranking journals (JIF-based ranking) from all Journal Citation Report categories

all journals. In contrast to the other characteristics of citation distribution which are similar across the cohorts of the current study, the percentage of CI receiving zero citations is rather heterogeneous: while the Top 3 journals from all JCR categories comprise only a median of 9% CI without citations in the JIF window, in the Med-R&E and Med-G&I categories, median 24% and 44% of papers have not been cited, respectively. Although further detailed analyses are need to prove this assumption, this difference could be explained by the Matthews-effect (Larivière and Gingras 2010), i.e. that the relatively high JIF of the Top 3 journals within each category tends to attract citations to all published items in these journals thus reducing the proportion of non-cited articles. As shown in Table 3, other studies reported values in the range between < 20 and 70% of CI (Weale et al. 2004; Asaad et al. 2019; Opthof et al. 2004; Lustosa et al. 2012; Bozzo et al. 2017). Weale and colleagues (Weale et al. 2004) discussed the percentage of articles per journal receiving no citations (rate of non-citation) as a possible alternative for measuring journal quality due to their observation that high-JIF journal have lower rates of non-citations. The main advantage could be that this approach might reduce the "temptation to use a journal's ranking to judge individual articles" (Weale et al. 2004). The present study confirms their findings, i.e. the negative correlation between non-citation rates and the journals' JIF is strong and highly significant: the Spearman rank correlation coefficient is almost identical with $-0.9581/-0.9830/-0.8607$ for the three cohorts in the present study compared to -0.854 and -0.924 in (Weale et al. 2004). Additionally, the current analysis clearly shows that below a threshold of e.g. JIF = 10 (~25 in the Top 3 cohort), the rate of non-citation dramatically increases (see Online resource 3-c), i.e. the rate of being not cited in these journals can readily be 50%. Therefore, especially worth considering in the lower segment of JIF-ranked journals, the general infeasibility of the JIF to predict individual article performance is underscored by the considerable high fraction of non-cited papers.

Although the summary of in Table 3 does not rise the claim for completeness and the included studies are rather heterogeneous with respect to the numbers of journals analysed (and how they were selected), the percentage of citable items contributing 50/90% to the total of the journals' citations is very similar across previous studies and our current results: between 15 and 25% of papers published in any journal analysed receive 50% of the citations to that journal.

This is further illustrated in Fig. 5: studies comprising larger sets of journals report between 12–20% of citable items per journal being responsible for 50% of all citations. Similarly homogeneous are the results on % citations received by the most frequently cited 50% of citable items: in published studies including our current results, usually 85–90% of all citations to a journal are generated by the best-cited 50% CI within that journal. It seems altogether that these indicators (%CI required for 50% citation and % citations generated by the best-cited 50% of papers) are quite robust and show no obvious dependency on the selection and number of journals and the date of analysis.

Correlation analyses as summarized in Table 2 and Online Resource 4 reveal a relationship between the various measures of unequal citation distribution with the JIF (mean citations): for all three investigated cohorts, journals with high JIFs have significantly i) higher percentages of CI contributing 50/90% of total citations, ii) lower numbers of non-cited CI, and iii) lower Gini coefficients. This could be interpreted as an indication that for high-JIF journals, the unequal distribution of citations is less pronounced than in low-JIF journals and the associated measures of inequality tend to be smaller. In further consequence, one could argue that the problem is less important for such high-JIF journals and the JIF would, therefore, adequately reflect those journals' high quality and impact. This argument supports the previously expressed opinion whether "in each speciality the best
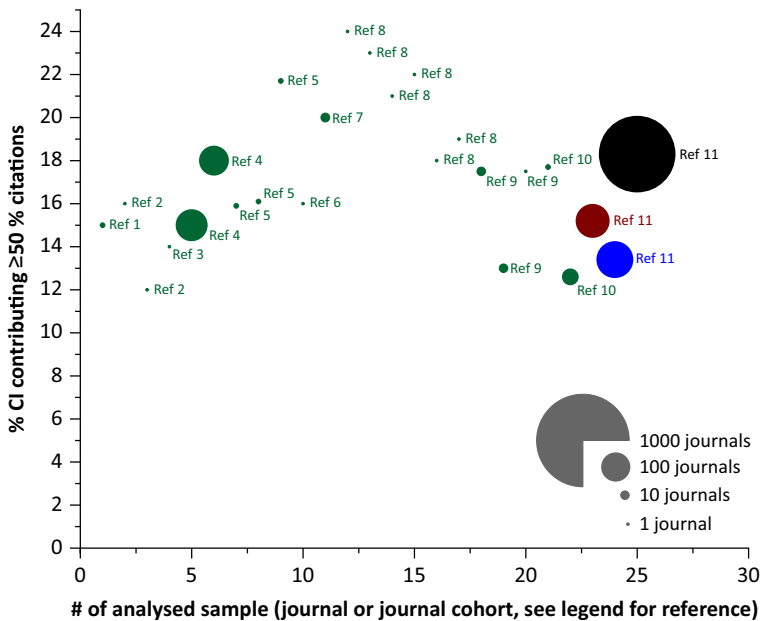
Fig. 5 Summary of previously published and current results on the %CI contributing ≥ 50% citations. The number of journals investigated in the individual studies is coded by the area of the circle. The label refers to the following studies: Ref 1 = (Seglen 1992), Ref 2 = (Opthof and Coronel 2002), Ref 3 = (Opthof et al. 2004), Ref 4 = (Weale et al. 2004), Ref 5 = (Falagas et al. 2010), Ref 6 = (Editor(s) 2011), Ref 7 = (Larivière et al. 2016), Ref 8 = (Zhang et al. 2017), Ref 9 = (Pang 2019), Ref 10 = (Asaad et al. 2019), Ref 11 = the current study (red = Medicine, Research & Experimental, blue = Medicine, General & Internal, black = the three highest ranking journals (JIF-based ranking) from all *Journal Citation Reports* categories). *Abbreviations*: *CI* citable items

journals are those (…) that have a high impact factor" and "the use of the impact factor as a measure of quality is widespread because it fits well with the opinion we have in each field of the best journals in our speciality" (Garfield 2006; Hoeffel 1998). While it is often indeed the high(est)-JIF journals where it is most difficult to have a manuscript accepted (Hoeffel 1998)—partly due to editorial policies aiming at attracting the most citable (trendiest, mainstream) articles (Falagas and Alexiou 2008; Taylor et al. 2008)—this argument, however, can be turned by saying "if you are a mature and active scholar in your field, you do not need the JIF (or any other metric) to know which journals are the best" (Browman and Stergiou 2008). In line with this, a reliable, robust and more meaningful metric of journal quality would be most important in the lower-JIF segment of (less well known, less prestigious) journals, as for those citation inequality is most pronounced. Due to the inability of the JIF to adequately represent the 'average' citation a journal receives (especially for lower JIF journals with highly skewed citation distribution), any JIF-based ranking (including three decimal digits) in this segment is quite meaningless. As recently shown (Koelblinger et al. 2019), journals publishing small numbers of papers show more pronounced JIF changes over time. Therefore, in addition to general citation skewness, in case of lower volume journals, the sample size (number of CI per journals) the JIF is based on further questions the significance of the JIF and the relevance of its temporal dynamics.

Due to the known skewness of citations, the median has been suggested as more appropriate to represent the central tendency of the citations distribution (e.g. (Editor(s) 2011; Opthof 2019; Pulverer 2013, 2015; Weale et al. 2004)). The current results indicate that this measure is considerably smaller than the mean (40–50% reduction of the numerical value) thus indicating the bias of the latter by highly-cited articles as reported previously (Colquhoun 2003; Larivière et al. 2016; Pulverer 2015; Seglen 1992). Additionally, when journals are ranked by median instead of mean (JIF), up to 40% of journals in the JCR category Med-G&I drop to a value of zero. These results are in line with Bozzo et al. who showed that for 74 orthopaedic journals, the median number of citations is zero for the majority of journals, i.e. 67 journals (90.5%) (Bozzo et al. 2017). Furthermore, the number of rankings is reduced to 13 instead of 135 or 159 rank positions for the categories Med-R&E and Med-G&I, respectively—resulting in individual journals at high median-based ranking positions and lower ranking positions comprising rather large numbers of journals. Similar observations were reported by (Pang 2019) who identified three broad groups which sufficiently rank veterinary journals and are separated by a median difference of one. As discussed by the creator of the JIF, E. Garfield (Garfield 2006), "The precision of impact factors is questionable, but reporting to 3 decimal places reduces the number of journals with the identical impact rank." Due to the small differences in JIF, "however, it matters very little whether, for example, the impact of JAMA is quoted as 24.8 rather than 24.831." (Garfield 2006). The current data support this statement: the median, supposed to better characterise journals in terms of an 'average' citation rate, results in a value of one or zero for virtually all journals in the Med-R&E and Med-G&I categories which have a JIF of 1–2 or below 1, respectively. The present study clearly confirms previous data showing that mean citations (JIF) are not suitable to represent the citation rate of a journal. While median citations are more appropriate for skewed citation distributions, their power to separate and rank journal in the lower segment of citation counts is quite limited.[2] Since any measure of central tendency captures only a small part of the information (Adler et al. 2009), it must be supplemented by information on the underlying distribution to adequately reflect the data; such information could be (1) the measures of inequality used here and in previous studies or, probably more familiar for most researchers, (2) e.g. interquartile ranges of citations. Altogether, as discussed by Adler at al. (Adler et al. 2009), in the current "culture of numbers" we should be aware of the illusory accuracy and seductive precision of crude statistics such as the JIF.

One limitation of the current study might be that instead of 'official' JIF values, manually calculated arithmetic mean values were used for those analyses involving the journals' JIF. As mentioned above, this approach was chosen in order to ensure comparability with the manually calculated median values. However, with this procedure, the already previously noted problem (e.g. (Glänzel and Moed 2002; Larivière et al. 2016)) with traceability of the JIF calculation became obvious (Online Resource 3): while most journals in all three investigated cohorts experience an approximately 5–10% value reduction by manual calculation (i.e. sum of citations to citable items as provided in the JCR database), extreme examples of journals show up to −30/−50/−80% lower values than the official JIFs in the cohorts Med-R&E, Med-G&I and Top 3, respectively. Interestingly, while all $n = 982$ journals in the current study's cohorts showed lower mean citation values than their JIFs due

---

[2] Although the publisher of the JCR, Clarivate Analytics (https://www.jcr.clarivate.com), has recently started to provide article and review median citation rates on the individual 'Journal Profile' pages, the median is not implemented in the journal list views and thus not available for sorting and ranking journals.

to higher numerators in the JIF equation compared to actual citations referring to CI only, previous data (Pang 2019) also report (minor) positive numerical changes when means were used instead of the JIF for individual journals. In contrast, for a selection of 30 most-cited cardiovascular journals, Opthof T. (Opthof 2019) demonstrated that the JIF of these journals is $15.0 \pm 3.5\%$ (mean $\pm$ s.e.m.) higher that a mean citation value considering citations to citable items in the numerator only. Irrespective of such contrasting results, the underlying problem consists in the JIF numerator containing citations for which no corresponding citable item is counted in the denominator (Larivière et al. 2016; Opthof 2019; Rossner et al. 2007). Although the classification of citable items in the JIF's denominator was claimed to be "accurate and consistent" by employees of the previous owner of the JCR database (McVeigh and Mann 2009), current and previous results (e.g. (Opthof 2019)) demonstrate a considerable effect of citations to non-citable items. This asymmetry provides room for negotiations between journals and the publisher of the JIF on whether which items should be counted towards the denominator (Editor(s) 2006; Rossner et al. 2007). In addition to the limited replicability of the JIF, the whole process of calculating JIFs and rating sciences by 'journal impact' has, therefore, been called "unscientific and arbitrary" and "unscientific, subjective, and secretive" (Editor(s) 2006) and "reflects the lack of transparency surrounding items included in the calculation, in contrast to the standards expected of published research" (Pang 2019). Noteworthy, the current study's results on citation inequality would be even more pronounced if the official, published JIF would have been used instead of the mean citations to citable items (used in this study).

Another limitation of the present study is the use of the Top 3 cohort of journals. While this cohort was intended to serve as comparison demonstrating the generality of the statements concerning citation inequality, such analyses could be expanded to larger and more representative journal selections in future studies covering more than the "three best journals" per category.

Conclusions

Referring to the initial question of the article's title, yes, it does matter: by analysing a large sample of medical and a selection of non-medical journals, the present analysis confirms a considerable inequality in citation distribution as illustrated by several quantitative measures—thus, disqualifying the JIF to adequately represent a journal's or a paper's citedness. While replacing the JIF (mean) by median citations does not necessarily overthrow journal rankings in the upper JIF segment, for lower-JIF journals, however, the 3-decimal-digit precision of JIF-based rankings is obviously meaningless when individual journals', articles' (or authors') citedness should be inferred. In further consequence, the current results provide additional up-to-date evidence for why assessing scientific quality or performance of individual authors or institutions by the JIF must be considered an inappropriate and mostly meaningless (mis-)use of this metric (Casadevall and Fang 2014; Garfield 1996; McKiernan et al. 2019; Simons 2008). In that context, instead of falling into the trap of a simple number pretending to represent a complex endeavour, multiple criteria need to be applied since science and research themselves have multiple dimensions and goals (Adler et al. 2009).

advice, and editing of the final draft. AT was involved in conceptualization, supervision of the study, and editing of the final draft.

Availability of data and materials Primary data this article is based on are not shared openly due to restricted, subscription-based access to the source database (JCR). The methods and procedures of data collection and analysis are fully disclosed in the Materials and methods section under "Data collection" and "Descriptive Analysis", respectively. The raw data are fully available to colleagues with an active institutional or personal subscription to the Clarivate Analytics JCR database (https://jcr.clarivate.com).

Code availability Not applicable.

Compliance with ethical standards

Conflicts of interest The authors declare that they have no conflict of interest.

References

Adler, R., Ewing, J., & Taylor, P. (2009). Citation statistics: A report from the International Mathematical Union (IMU) in Cooperation with the International Council of Industrial and Applied Mathematics (ICIAM) and the Institute of Mathematical Statistics (IMS). *Statistical Science, 24*(1), 1–14.

Albarrán, P., Crespo, J. A., Ortuno, I., & Ruiz-Castillo, J. (2011). The skewness of science in 219 sub-fields and a number of aggregates. *Scientometrics, 88,* 385–397.

Asaad, M., Kallarackal, A. P., Meaike, J., Rajesh, A., de Azevedo, R. U., & Tran, N. V. (2019). Citation Skew in Plastic Surgery Journals: Does the journal impact factor predict individual article citation rate? *Aesthetic Surgery Journal.* https://doi.org/10.1093/asj/sjz336.

Bornmann, L., & Leydesdorff, L. (2017). Skewness of citation impact data and covariates of citation distributions: A large-scale empirical analysis based on Web of Science data. *Journal of Informetrics, 11,* 164–175. https://doi.org/10.1016/j.joi.2016.12.001.

Bozzo, A., Oitment, C., Evaniew, N., & Ghert, M. (2017). The Journal Impact Factor of Orthopaedic Journals Does not Predict Individual Paper Citation Rate. *Journal of the American Academy of Orthopaedic Surgeons: Global Research and Reviews, 1*(2), e007. https://doi.org/10.5435/JAAOSGlobal-D-17-00007.

Browman, H., & Stergiou, K. (2008). Factors and indices are one thing, deciding who is scholarly, why they are scholarly, and the relative value of their scholarship is something else entirely. *Ethics in Science and Environmental Politics, 8,* 1–3.

Campbell, P. (2008). Escape from the impact factor. *Ethics in Science and Environmental Politics, 8,* 5–7.

Casadevall, A., & Fang, F. C. (2014). Causes for the persistence of impact factor mania. *mBio, 5*(2), e00064-00014. https://doi.org/10.1128/mBio.00064-14.

Colquhoun, D. (2003). Challenging the tyranny of impact factors. *Nature, 423*(6939), 479. https://doi.org/10.1038/423479a.

De Maio, F. G. (2007). Income inequality measures. *Journal of Epidemiology and Community Health, 61*(10), 849–852. https://doi.org/10.1136/jech.2006.052969.

Editor(s) (2006). The impact factor game. It is time to find a better way to assess the scientific literature. *PLoS Medicine 3*(6), 291, https://doi.org/10.1371/journal.pmed.0030291.

Editor(s) (2011). Dissecting our impact factor. *Nature Materials, 10*(9), 645, https://doi.org/10.1038/nmat3114.

Falagas, M. E., & Alexiou, V. G. (2008). The top-ten in journal impact factor manipulation. *Archivum Immunolgiae et Therapiae Experimentalis, 56*(4), 223–226. https://doi.org/10.1007/s00005-008-0024-5.

Falagas, M. E., Kouranos, V. D., Michalopoulos, A., Rodopoulou, S. P., Batsiou, M. A., & Karageorgopoulos, D. E. (2010). Comparison of the distribution of citations received by articles published in high, moderate, and low impact factor journals in clinical medicine. *Internal Medicine Journal, 40*(8), 587–591. https://doi.org/10.1111/j.1445-5994.2010.02247.x.

Garfield, E. (1955). Citation indexes for science; a new dimension in documentation through association of ideas. *Science, 122*(3159), 108–111.

Garfield, E. (1996). How can impact factors be improved? *BMJ, 313*(7054), 411–413. https://doi.org/10.1136/bmj.313.7054.411.

Garfield, E. (2006). The history and meaning of the journal impact factor. *JAMA, 295*(1), 90–93. https://doi.org/10.1001/jama.295.1.90.

Gini, C. (2005). On the measurement of concentration and variability of characters (translation by Giovanni Maria Giorgi). *METRON - International Journal of Statistics, LXIII*(1), 3–38.

Glänzel, W., & Moed, H. F. (2002). Journal impact measures in bibliometric research. *Scientometrics, 53*(2), 171–193.

Gross, P. L. K., & Gross, E. M. (1927). College Libraries and Chemical Education. *Science, 66*(1713), 385–389. https://doi.org/10.1126/science.66.1713.385.

Henderson, A. R. (2006). Testing experimental data for univariate normality. *Clinica Chimica Acta, 366*(1–2), 112–129. https://doi.org/10.1016/j.cca.2005.11.007.

Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature, 520*(7548), 429–431. https://doi.org/10.1038/520429a.

Hoeffel, C. (1998). Journal impact factors. *Allergy, 53*(12), 1225. https://doi.org/10.1111/j.1398-9995.1998.tb03848.x.

Koelblinger, D., Zimmermann, G., Weineck, S. B., & Kiesslich, T. (2019). Size matters! Association between journal size and longitudinal variability of the Journal Impact Factor. *PLoS ONE, 14*(11), e0225360. https://doi.org/10.1371/journal.pone.0225360.

Larivière, V., & Gingras, Y. (2010). The Impact Factor's Matthew effect: A natural experiment in bibliometrics. *Journal of the American Society for Information Science, 61*(2), 424–427.

Larivière, V., Kiermer, V., MacCallum, C. J., McNutt, M., Patterson, M., Pulverer, B., et al. (2016). A simple proposal for the publication of journal citation distributions. *bioRxiv*. https://doi.org/10.1101/062109.

Larivière, V., & Sugimoto, C. R. (2019). The journal impact factor: A brief history, critique, and discussion of adverse effects. In W. Glänzel, H. F. Moed, U. Schmoch, & M. Thelwall (Eds.), *Springer handbook of science and technology indicators*. New York: Springer.

Lustosa, L. A., Chalco, M. E., Borba Cde, M., Higa, A. E., & Almeida, R. M. (2012). Citation distribution profile in Brazilian journals of general medicine. *Sao Paulo Medical Journal, 130*(5), 314–317. https://doi.org/10.1590/s1516-31802012000500008.

McKiernan, E. C., Schimanski, L. A., Munoz Nieves, C., Matthias, L., Niles, M. T., & Alperin, J. P. (2019). Use of the Journal Impact Factor in academic review, promotion, and tenure evaluations. *eLife*. https://doi.org/10.7554/eLife.47338.

McVeigh, M. E., & Mann, S. J. (2009). The journal impact factor denominator: defining citable (counted) items. *JAMA, 302*(10), 1107–1109. https://doi.org/10.1001/jama.2009.1301.

Opthof, T. (1997). Sense and nonsense about the impact factor. *Cardiovascular Research, 33*(1), 1–7. https://doi.org/10.1016/s0008-6363(96)00215-5.

Opthof, T. (2019). Comparison of the Impact Factors of the most-cited Cardiovascular Journals. *Circulation Research, 124*(12), 1718–1724. https://doi.org/10.1161/CIRCRESAHA.119.315249.

Opthof, T., & Coronel, R. (2002). The impact factor of leading cardiovascular journals: Where is your paper best cited? *Netherlands Heart Journal, 10*(4), 198–202.

Opthof, T., Coronel, R., & Piper, H. M. (2004). Impact factors: No totum pro parte by skewness of citation. *Cardiovascular Research, 61*(2), 201–203. https://doi.org/10.1016/j.cardiores.2003.11.023.

Opthof, T., & Wilde, A. A. (2009). The Hirsch-index: A simple, new tool for the assessment of scientific output of individual scientists: The case of Dutch professors in clinical cardiology. *Netherlands Heart Journal, 17*(4), 145–154. https://doi.org/10.1007/BF03086237.

Pang, D. S. J. (2019). Misconceptions surrounding the relationship between journal impact factor and cita-
    tion distribution in veterinary medicine. *Veterinary Anaesthesia and Analgesia, 46*(2), 163–172. https
    ://doi.org/10.1016/j.vaa.2018.11.004.
Pulverer, B. (2013). Impact fact-or fiction? *EMBO Journal, 32*(12), 1651–1652. https://doi.org/10.1038/
    emboj.2013.126.
Pulverer, B. (2015). Dora the brave. *EMBO Journal, 34*(12), 1601–1602. https://doi.org/10.15252/
    embj.201570010.
Rossner, M., Van Epps, H., & Hill, E. (2007). Show me the data. *Journal of Cell Biology, 179*(6), 1091–
    1092. https://doi.org/10.1083/jcb.200711140.
Rostami-Hodjegan, A., & Tucker, G. T. (2001). Journal impact factors: A 'bioequivalence' issue? *British
    Journal of Clinical Pharmacology, 51*(2), 111–117. https://doi.org/10.1111/j.1365-2125.2001.01349
    .x.
Seglen, P. O. (1989). From bad to worse: Evaluation by Journal Impact. *Trends in Biochemical Sciences,
    14*(8), 326–327. https://doi.org/10.1016/0968-0004(89)90163-1.
Seglen, P. O. (1992). The skewness of science. *Journal of the American Society for Information Science and
    Technology, 43*(9), 628–638.
Seglen, P. O. (1998). Citation rates and journal impact factors are not suitable for evaluation of research.
    Acta Orthopaedica Scandinavica, 69(3), 224–229.
Simons, K. (2008). The misused impact factor. *Science, 322*(5899), 165. https://doi.org/10.1126/scien
    ce.1165316.
Taylor, M., Perakakis, P., & Trachana, V. (2008). The siege of science. *Ethics in Science and Environmental
    Politics, 8,* 17–40.
Weale, A. R., Bailey, M., & Lear, P. A. (2004). The level of non-citation of articles within a journal as a
    measure of quality: A comparison to the impact factor. *BMC Medical Research Methodology, 4,* 14.
    https://doi.org/10.1186/1471-2288-4-14.
Wilsdon, J., Allen, L., Belfiore, E., Campbell, P., Curry, S., Hill, S., et al. (2015). *The Metric Tide: Report of
    the independent review of the role of metrics in research assessment and management*. HEFCE. https://
    doi.org/10.13140/RG.2.1.4929.1363.
Zhang, L., Rousseau, R., & Sivertsen, G. (2017). Science deserves to be judged by its contents, not by
    its wrapping: Revisiting Seglen's work on journal impact and research evaluation. *PLoS ONE, 12*(3),
    e0174205. https://doi.org/10.1371/journal.pone.0174205.

Affiliations

Tobias Kiesslich[1,2] · **Marlena Beyreis**[1] · **Georg Zimmermann**[3,4,5] ·
Andreas Traweger[6,7]

✉ Tobias Kiesslich
    tobias.kiesslich@pmu.ac.at

1   Institute of Physiology and Pathophysiology, Paracelsus Medical University, Strubergasse 21,
    5020 Salzburg, Austria

2   Department of Internal Medicine I, Paracelsus Medical University/Salzburger Landeskliniken,
    Muellner Hauptstrasse 48, 5020 Salzburg, Austria

3   Team Biostatistics and Big Medical Data, IDA Lab Salzburg, Paracelsus Medical University,
    Strubergasse 16, 5020 Salzburg, Austria

4   University Clinic of Neurology, Christian Doppler Medical Centre, Paracelsus Medical University,
    Ignaz Harrer Straße 79, 5020 Salzburg, Austria

5   Department of Mathematics, Paris-Lodron University of Salzburg, Hellbrunnerstrasse 34,
    5020 Salzburg, Austria

6   Institute of Tendon and Bone Regeneration, Spinal Cord Injury and Tissue Regeneration Center
    Salzburg, Paracelsus Medical University, Strubergasse 21, 5020 Salzburg, Austria

7   Austrian Cluster for Tissue Regeneration, Donaueschingenstrasse 13, 1200 Vienna, Austria