Measuring an individual researcher's impact – new directions and challenges

Morwenna Senior[1], Seena Fazel[1]

Department of Psychiatry, University of Oxford, Oxford, UK

Correspondence to Seena Fazel, Department of Psychiatry, Warneford Hospital, Warneford Lane, Oxford, OX3 7JX; seena.fazel@psych.ox.ac.uk

This article has been accepted for publication in *Evidence-Based Mental Health*, 2019 following peer review. The Version of Record can be accessed online, doi: 10.1136/ebmental-2019-300122.

Abstract
Metrics which quantify the impact of a scientist are increasingly incorporated into decisions about how to rate and fund individuals and institutions. Several commonly used metrics, based on journal impact factors and citation counts, have been criticised as they do not reliably predict real-world impact, are highly variable between fields, and are vulnerable to gaming. Bibliometrics have been incorporated into systems of research assessment but these may create flawed incentives, failing to reward research that is validated, reproducible and with wider impacts. A recent proposal for a new standardised citation metric based on a composite score of 6 measures has led to an online database of the 100,000 most highly cited scientists in all fields. In this perspective article, we provide an overview and evaluation of this new citation metric as it applies to mental health research. We provide a summary of its findings for psychiatry and psychology, including clustering in certain countries and institutions, and outline some implications for mental health research. We discuss strengths and limitations of this new metric, and how further refinements could align impact metrics more closely with wider goals of scientific research.

KEYWORDS: bibliometrics, impact metric, journal impact factor, citation metric, altmetrics

Introduction

Estimating the impact of a scientist's research can contribute to individual-level decisions, such as funding and promotion, and the evaluation of institutions as they compete for public funding and seek to enhance their reputation and attract students.[1] Increasingly, quantitative metrics are used in these decisions as one way to include more objectivity, accountability and transparency in decision-making. In so doing, metrics complement alternative approaches that typically use a system of peer or expert review, as used in the UK's Research Excellence Framework. Although peer review is associated with substantial resource implications and possible risk of bias, many have argued it is necessary as metric-based approaches used in isolation may not provide meaningful estimates of research impact and are highly variable between fields.

The need for improving these metrics is not only a question of optimising decision-making processes. Research assessment creates a system of rewards and incentives that can shape what is researched and communicated.[2,3] At their worst, poor systems can promote low quality research, undermine scientific integrity and hinder translation of research findings to clinical practice.[2] Alternatively, if fairly determined and used in responsible ways, metrics have the potential to promote research relevant to clinical practice, replicable and engaging to the wider public.

In this perspective article, we summarise and evaluate a new standardised citation metric for individual scientists proposed by Ioannidis and colleagues as it applies to mental health research and discuss the wider implications of metrics in research assessment.

Current metrics

Currently some metrics are used in research assessment, but they all have significant limitations. The Journal Impact Factor is not always strongly correlated with research quality and heavily influenced by a high number of citations to a minority of publications within a particular journal.[4,5] Simple citation counts are also inadequate, as they do not account for an individual author's contribution.[6] Genetic consortia provide a clear example of this where an author may contribute information on a few individuals, and this is used in several publications that report on different outcomes or updated analyses. This is complicated by differences in how institutions and academic cultures determine what constitutes authorship. Consistent with this, 'hyperprolific' authors, who have published >70 papers in a year, cluster in certain institutions and countries.[7] Such authorship practices may have been encouraged by research assessment that privileges authorship, making the number of publications the most valuable currency by which a scientist's output is rated. An additional challenge to any metric is the extent to which it can be gamed with tactical citation practices or collaborative citation schemes.[1]

A new impact metric

A recent paper published in *PLoS Biology* by Ioannidis and colleagues proposes a new standardised citation metric to assess individual scientists.[6,8] This composite score aims to address some of the problems associated with commonly used metrics by accounting for differences in author contributions and reducing risk of questionable citation practices. The unit of assessment is an individual scientist. The score is based on routinely collected citation and authorship data from Scopus. It is a composite of six citation metrics: total citations; Hirsch h-index; Schreiber hm-index (which adjusts for co-authorship); number of citations to papers where the scientist is the only author; number of citations to papers where the scientist is the first or only author; and number of citations to papers where the scientist is the first, only, or last author.[6] These six metrics are equally weighted and combined to create the c-score out of a total score of 6. This composite score has been calculated in two freely available databases on Mendeley of the 100,000 top-cited scientists across all fields, which include each of the components of the composite score with and without self-citations, and the country and institution for each scientist. The two databases differ in the timescale of the citation metrics reported: one is based on citations in all years providing an overview of career impact; the second provides each of the component metrics and the composite score based on citations in 2017, for example the 1-year h-index is based on citations in 2017. We have focused on the 2017 database as it is less influenced by long publication histories and author age. One problem is that the institutional affiliation may be unavailable or misclassified in this database for some authors, particularly those with multiple affiliations (available at https://data.mendeley.com/datasets/btchxktzyw/1).

From this database, we extracted the 7116 most highly cited researchers publishing in psychiatry, psychology and related fields according to their composite score. This version uses citations in the year 2017 to papers published since 1960 (see Supplementary file 1 ranked by composite score, and Supplementary file 2 ordered by institution). These scientists were selected if they were categorised as publishing in psychiatry or psychology-related fields as primary or secondary fields. This list is therefore inclusive, incorporating for example experimental/developmental psychology, those for whom psychiatry/psychology is the second-ranked field for publications, and many individuals working primarily in neurology and neuroscience. However, we have excluded 'social' and 'comparative' psychology as they are not directly relevant to mental health research. Table 1 shows citation metrics for the 20 highest-ranked of these researchers in 2017 (ranked according to their composite score removing self-citations). Table 1 also contains the 5 next most highly-ranked scientists with psychiatry as a first of second field, giving a narrower definition of the field (see extended list of scientists working psychiatry in supplementary file 3).

To examine how the c-score differs from other common metrics for individual authors, we compared ranking based on composite-score and 1-year h-index for the top ten highest-cited scientists in psychiatry and psychology-related fields according to the composite score (Supplementary file 7). For two scientists, the h-index resulted in much lower rankings: from 7 (c-score) to 260 (1-year h-index), and from 9 (c-score) to 119 (1-year h-index). These differences are driven by large number of citations to papers as single/first author for these scientists, which have contributed to the relatively high composite score compared to the h-index which does not account for co-authorship.

Table 2 shows the country of affiliation for the top-cited scientists in psychiatry, psychology and related fields. 83% of the highly ranked scientists (5528 out of 6664 with a known affiliation) were based in five countries: USA, UK, Canada, Germany, and Australia (supplementary file 4). These five countries may be distinguished by well-funded research networks, but the relatively large number of highly cited scientists may also be influenced by publication and citation practices within these countries.[9] The top-cited mental health scientists had 819 institutional affiliations, and Harvard University had the largest number of affiliated scientists (n=177) followed by the University of Toronto (n=109), King's College London (n=108), University of California at Los Angeles (n=106) and Columbia University (n=100) (Supplementary file 5).

We have also cross-referenced this database with a list of 'hyperprolific authors' (defined as publishing >70 articles in a year), which was outlined in another paper in *Nature* 2019 (Supplementary file 6).[7] 14 authors appear both among highly cited scientists in psychiatry/psychology and the hyperprolific list (out of a total of 140 hyperprolific authors in biology/health sciences).

Strengths and limitations

This new approach to measuring impact (and the online database) is an advance as it attempts to balance citations with authorship, and provides one way to compare citation practices and metrics between individuals and institutions. At the same time, although the database allows comparison between fields, this needs to be done with caution. The equal weighting of the six citation metrics that make up the composite score may favour some subfields. One example is that single author papers, which are one of the six metrics, are rare in some disciplines.

In addition, the classification of researchers into subject fields is problematic for identifying comparable researchers. Categorisation is based on classification of journals in which the authors publish. As a result, authors who collaborate across multiple fields may be more likely to be misclassified, especially if they publish in more general journals, although most researchers are unlikely to have all their research output in such venues. More problematic is that this type of categorization by journal means that the author listing is over-inclusive, in particular for neurologists, social psychologists, statisticians and some others. The Relative Citation Rate (RCR) is an example of a recently developed metric which defines scientific fields using a different approach. This article-based metric measures citation rates relative to a co-citation network of papers which are cited alongside the article of interest.[10] Thus, the RCR employs a user-driven, flexible definition of the field. Compared to the field definitions provided by Ioannidis et al, the co-citation network may better reflect interdisciplinary fields, however it has been criticised for potentially penalising papers which are cited outside of their subject area.[11] The approach of the RCR is to normalise to a field which is specific to each paper and not readily scrutinised, whereas the proposed c-score is more transparent and can be adjusted according to the needs of the user. The author-level RCR metric that can be generated through the NIHR's icite interface,[12] does not account for co-authorship unless users manually select only a subset of an author's papers. This differs from the approach of the c score which incorporates the Hm index (accounting for the number of co-authors) and citations to papers as first/last/only author.

A further limitation is that the composite citation score is higher for scientists with a long publication history, and thus who are older. As such, the database contains few individuals who are early in their research career and does not capture individual research trajectories. Incorporating only citations to papers published in recent years might avoid this limitation, but is problematic because recent publications may not yet have accumulated citations.[6]

In addition, how author contribution is characterised can be criticized. Using single/first/last author status as a proxy for author contribution partly depends on authorship practices.[13] As such, clarification of authorship contributions, if done precisely, provides an important complement to this. However, how and whether this can be metricized needs further thought. One approach could be the use of structured author taxonomies which categorize types of contributions, such as the CRediT taxonomy used by PLoS journals.[14] To incorporate these taxonomies into impact metrics, they would ideally need to be used routinely by journals and published.

Impact metrics in context

The use of impact metrics within mental health needs more debate. A range of approaches should be considered. For applied research, a change in clinical practice may be the most important outcome.[15] This might be reflected in the uptake of research findings in clinical guidelines, which may not be reflected in citation counts. Other individual contributions are less easily counted, such as mentoring junior colleagues, development and maintenance of software and research infrastructure (including databases), and leadership roles within departments or collaborative projects.[2]

Measures of wider public interest in an individual's research could also be considered.[16] However, their utility should be appraised carefully as they may not measure forms of attention which reflect valuable impact or research quality. Some alternative metrics aim to incorporate news and social media coverage. One example is the number and extent of twitter mentions, which contribute to the Altmetric attention score. Case studies of highly tweeted papers have highlighted examples where large numbers of mentions are driven by one or two accounts which focus repeatedly on a single topic, or patterns of repetitive tweeting likely to come from automated 'bots'.[17] Another example showed that allegations of scientific misconduct drove high levels of twitter attention.[18] In addition, twitter counts are poorly correlated with judgements of research quality made by experts or peers.[19,20]

Several suggestions have been made for 'best practice' in research assessment which may guide decisions about the use of impact metrics in mental health.[1,4,21] Among these is the need to first establish clear, context-specific priorities for research. Changes in research assessment processes (including the use of novel metrics) should be assessed with outcomes relevant to these priorities. One such priority is the need to reward replication or validation studies. Psychiatry and psychology are among several fields where replication attempts have been disappointing.[22,23]

Another priority for mental health might be rewarding interdisciplinarity, which may not be adequately acknowledged by current metrics. One example is reflected in priorities for mental health research that were recently set out by the UK's Department of Health.[24] This report emphasizes the need to research mental health interventions in non-healthcare settings such as schools, prisons, and care homes. Further cross-disciplinary work will be necessary to meet this goal. Each of the fields involved may have different approaches to citation and authorship practices which could affect the comparability of impact metrics.

In conclusion, Ioannidis and colleagues have produced a composite metric that clearly improves on simple publication or citation counts by accounting for authorship contribution and providing tools to examine citation practices. In this paper, we have provided a ranking of individuals in psychiatry, psychology and related fields based on 2017 metrics, and provided an overview of the countries and institutions with the highest ranked scientists. We have argued that impact metrics need interpretation in a context-specific manner, designing and assessing processes which align to the wider priorities of the particular field. This is also the case for journals [5] and individual publications [16], and this new scores provides a way to evaluate individual scientists. Further refinements to this approach might assist in developing more scalable and transparent ways to assess individual research contributions.

Competing interests: none declared.

Funding: There was no specific funding.

Table 1: Citation metrics based on 2017 citations for the 20 top-ranked scientists publishing in psychiatry, psychology and related fields, and additional top 5 scientists publishing in psychiatry, ranked according to composite score (c score).

Author name	Institute	Country	Papers 1960 – 2017 (n)	c score	total citations	1-year h index 2017	1-year hm index	Citations to single-author papers	Citations to papers as single/first author	Citations to papers as first/single/last author	% of citations which are self-citations
Top 20 most highly-cited scientists publishing in fields related to psychiatry and psychology											
Kessler, Ronald C.	Harvard University	USA	645	5.22	12838	50	23.0	354	6478	9848	5.15%
Friston, Karl J.			701	5.15	8444	37	24.6	1167	3193	6170	10.71%
Hinton, Geoffrey E.	Google Inc.	USA	152	5.11	16476	34	17.6	549	3625	14124	0.14%
Smith, Stephen M.	University of Oxford	GBR	206	4.97	9239	45	14.4	615	3002	6098	4.15%
Gross, James J.	Stanford University	USA	280	4.89	4774	33	22.0	877	1574	4062	5.52%
Raichle, Marcus E.	Washington University St. Louis	USA	237	4.68	5112	34	17.5	314	1155	3691	4.16%
Baddeley, Alan	University of York	GBR	213	4.66	2818	20	15.4	1304	2067	2577	6.19%
Buckner, Randy L.	Harvard University	USA	212	4.64	5899	36	15.8	179	1507	3482	10.16%
Beck, Aaron T.	University of Pennsylvania	USA	248	4.62	4215	23	15.3	222	3221	3932	1.29%
Moffitt, Terrie E.	Duke University	USA	315	4.61	7303	32	15.2	361	1020	2099	6.28%
Fischl, Bruce	Harvard University	USA	215	4.61	5785	34	11.6	284	1890	3220	5.90%
Sporns, Olaf	Indiana University Bloomington	USA	138	4.60	4364	29	17.9	321	858	3430	4.99%
Tversky, Amos	Stanford University	USA	69	4.58	3412	24	15	277	2561	3387	0.15%
Nestler, Eric J.	Mount Sinai School of Medicine	USA	427	4.58	4724	32	16.8	277	676	3496	9.64%
Rutter, Michael	Medical Research Council	GBR	369	4.56	3417	24	15.6	682	1045	2197	4.39%
Volkow, Nora D.	National Institute on Alcohol Abuse and Alcoholism	USA	524	4.55	5095	32	17.9	45	2442	4128	6.67%
Steinberg, Laurence	Temple University	USA	180	4.54	2607	25	16.8	577	1172	1850	5.71%
Frith, Chris D.			409	4.52	4845	31	19.6	164	587	2978	3.60%
Koob, George F.	National Institute on Drug Abuse	USA	519	4.52	3346	23	16.6	350	1259	2423	13.11%
Nolen-Hoeksema, Susan	Yale University	USA	108	4.51	2368	25	16.9	445	1189	1919	1.66%
Additional 5 most highly-cited authors publishing specifically in psychiatry-related journals											
Kendler, Kenneth S.	Virginia Commonwealth University	USA	761	4.49	4824	26	13.7	182	1503	2714	18.33%
Cohen, Jacob	New York University	USA	28	4.47	2706	11	8.9	2463	2474	2607	0.00%
Robbins, Trevor W.	University of Cambridge	GBR	698	4.47	6077	30	18.7	111	442	3449	11.19%
Insel, Thomas R.			191	4.42	2587	21	14.6	405	1149	1864	2.85%
Corrigan, Patrick W.	Illinois Institute of Technology	USA	236	4.31	1972	20	13.4	267	1283	1584	9.71%

Note: All citation metrics based on citations in the year 2017. h score = Hirsch h index based on citations in 1 year, hm score = Schreiber Hm index based on citations for 1 year (which partially adjusts for co-authorship). Extracted from the database linked to the Ioannidis paper [8]. Other scientists publishing in psychiatry/psychology, and full metrics including self-citations are listed by composite score in Supplementary file 1, and by institution in supplementary file 2. A list of scientists publishing in psychiatry-related journals is in supplementary file 3, ranked by composite score.

Table 2: Country-level data on top-cited scientists in psychiatry, psychology and related fields.

Country	Top-cited scientists in psychiatry and psychology n (%)	Institutions with top-cited scientists in psychiatry and psychology affiliated (n)
USA	3638 (54.6)	316
Great Britain	812 (12.2)	84
Canada	428 (6.4)	39
Germany	399 (6.0)	66
Australia	251 (3.8)	33
Netherlands	230 (3.5)	16
Italy	110 (1.7)	35
Switzerland	102 (1.5)	10
Sweden	75 (1.1)	10
France	69 (1.0)	31
Belgium	68 (1.0)	7
Israel	67 (1.0)	8
Spain, Japan, New Zealand, Denmark, Norway, Finland	31 - 39 (0.5 – 0.99)	3 - 19

Note: Percentages are % of scientists with a known country affiliation. Table includes countries with >0.5% of scientists in the list of top-cited scientists in psychiatry, psychology and related fields. Excludes 452 scientists for whom country/institution information was not available. An extended table showing where all countries were affiliated is in supplementary file 4.

References

1 Wilsdon J, Allen L, Belfiore E, et al. *The metric tide : report of the Independent Review of the Role of Metrics in Research Assessment and Management*. HEFCE. 2015. doi:10.13140/RG.2.1.4929.1363

2 Moher D, Naudet F, Cristea IA, et al. Assessing scientists for hiring, promotion, and tenure. *PLoS Biol* 2018;**16**:e2004089. doi:10.1371/journal.pbio.2004089

3 Higginson AD, Munafò MR. Current Incentives for Scientists Lead to Underpowered Studies with Erroneous Conclusions. *PLoS Biol* 2016;**14**:e2000995. doi:10.1371/journal.pbio.2000995

4 American Society for Cell Biology. DORA. *The Declaration On Research Assessment*. 2012. sfdora.org/ (accessed Sep 2019).

5 Fazel S, Lamsma J. Beyond the impact factor? *Evid Based Ment Heal* 2015;**18**:33–5. doi:10.1136/eb-2015-102087

6 Ioannidis JPA, Klavans R, Boyack KW. Multiple Citation Indicators and Their Composite across Scientific Disciplines. *PLoS Biol* 2016;**14**:e1002501. doi:10.1371/journal.pbio.1002501

7 Ioannidis JPA, Klavans R, Boyack KW. Thousands of scientists publish a paper every five days. *Nature* 2018;**561**:167–9. doi:10.1038/d41586-018-06185-8

8 Ioannidis JPA, Baas J, Klavans R, et al. A standardized citation metrics author database annotated for scientific field. *PLOS Biol* 2019;**17**:e3000384. doi:10.1371/journal.pbio.3000384

9 Papatheodorou SI, Trikalinos TA, Ioannidis JPA. Inflated numbers of authors over time have not been just due to increasing research complexity. *J Clin Epidemiol* 2008;**61**:546–51. doi:10.1016/j.jclinepi.2007.07.017

10 Hutchins BI, Yuan X, Anderson JM, et al. Relative Citation Ratio (RCR): A New Metric That Uses Citation Rates to Measure Influence at the Article Level. *PLOS Biol* 2016;**14**:e1002541.

11 Waltman L. NIH's new citation metric: A step forward in quantifying scientific impact? 2015. https://www.cwts.nl/blog?article=n-q2u294 (accessed Nov 2019).

12 NIH. iCite | NIH Office of Portfolio Analysis. https://icite.od.nih.gov/analysis (accessed Nov 2019).

13 Sauermann H, Haeussler C. Authorship and contribution disclosures. *Sci Adv* 2017;**3**:e1700404. doi:10.1126/sciadv.1700404

14 Allen L, O'Connell A, Kiermer V. How can we ensure visibility and diversity in research contributions? How the Contributor Role Taxonomy (CRediT) is helping the shift from authorship to contributorship. *Learn. Publ.* 2019;**32**:71–4. doi:10.1002/leap.1210

15 Waddell C. So much research evidence, so little dissemination and uptake: mixing the useful with the pleasing. *Evid Based Ment Health* 2001;**4**:3–5. doi:10.1136/ebmh.4.1.3

16 Fazel S, Wolf A. What is the impact of a research publication? Evid Based Ment Health 2017;20:33–4. doi:10.1136/eb-2017-102668

17 Robinson-Garcia N, Costas R, Isett K, et al. The unbearable emptiness of tweeting—About journal articles. *PLoS One* 2017;**12**:e0183551. doi:10.1371/journal.pone.0183551

18 Bornmann L, Haunschild R. Allegation of scientific misconduct increases Twitter attention. *Scientometrics.* 2018;**115**:1097–100. doi:10.1007/s11192-018-2698-6

19 Bornmann L, Haunschild R. Do altmetrics correlate with the quality of papers? A large-scale empirical study based on F1000Prime data. *PLoS One* 2018;**13**:e0197133. doi:10.1371/journal.pone.0197133

20 Higher Education Funding Council for Europe. *Correlation analysis of REF2014 scores and metrics Supplementary Report II to the Independent Review*. HEFCE. 2015. doi:10.13140/RG.2.1.3362.4162

21 Hicks D, Wouters P, Waltman L, et al. Bibliometrics: The Leiden Manifesto for research metrics. *Nature*. 2015;**520**:429–31. doi:10.1038/520429a

22 Tajika A, Ogawa Y, Takeshima N, et al. Replication and contradiction of highly cited research papers in psychiatry: 10-year follow-up. *Br. J. Psychiatry.* 2015;**207**:357–62. doi:10.1192/bjp.bp.113.143701

23 Leichsenring F, Abbass A, Hilsenroth MJ, et al. Biases in research: Risk factors for non-replicability in psychotherapy and pharmacotherapy research. *Psychol. Med.* 2017;**47**:1000–11. doi:10.1017/S003329171600324X

24 United Kingdom Department of Health. *A Framework for mental health research.* Department of Health and Social Care. 2017:53. www.gov.uk/government/publications/a-framework-for-mental-health-research (accessed Sep 2019)