The data and analysis underlying NIH's decision to cap research support lacked rigor and transparency: a commentary

A. Cecile J.W. Janssens[1], Gary W. Miller[2], K.M. Venkat Narayan[1,3]

[1] Department of Epidemiology, [2] Department of Environmental Health, [3] Hubert Department of Global Health, Rollins School of Public Health, Emory University, Atlanta, USA.

Correspondence to: Professor A. Cecile J. W. Janssens PhD, Department of Epidemiology, Rollins School of Public Health, Emory University, 1518 Clifton Road NE, Atlanta, Georgia 30322, USA. E-mail: cecile.janssens@emory.edu, Telephone: +1 404 727 6307, Fax: +1 404 727 8737.

Abstract

The US National Institutes of Health (NIH) recently announced that they would limit the number of grants per scientist and redistribute their funds across a larger group of researchers. The policy was withdrawn a month later after criticism from the scientific community. Even so, the basis of this defunct policy was flawed and it merits further examination. The amount of grant support would have been quantified using a new metric, the Grant Support Index (GSI), and limited to a maximum of 21 points, the equivalent of three R01 grants. This threshold was decided based upon analysis of a new metric of scientific output, the annual weighted Relative Citation Ratio, which showed a pattern of diminishing returns at higher values of the GSI. In this commentary, we discuss several concerns about the validity of the two metrics and the quality of the data that the NIH had used to set the grant threshold. These concerns would have warranted a re-analysis of new data to confirm the legitimacy of the GSI threshold. Data-driven policies that affect the careers of scientists should be justified by nothing less than a rigorous analysis of high-quality data.

In May 2017, the US National Institutes of Health (NIH) announced that they would place a cap on the number of grants per scientist to redistribute funds across a larger group of researchers in an attempt to optimize output and maximize impact.[1 2] The proposal received immediate criticism from the scientific community. Researchers disputed the point values of the Grant Support Index (GSI), worried that the point system would discourage the application of training and collaborative grants, and urged the NIH to find other ways to support more researchers.[3] Within a month, the NIH withdrew the proposal and replaced it with the Next Generation Researchers Initiative, which aims to bolster the funding of early-stage and mid-career researchers without capping the number of grants for others.[4]

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3106v1 | CC BY 4.0 Open Access | rec: 25 Jul 2017, publ: 25 Jul 2017

The initial policy to limit the number of grants per researcher was put in place because the NIH had suggested that the output of investigators with multiple grants, on average, did not increase proportionately with their amount of funding compared to those with fewer grants.[1] The NIH had analyzed data on grant output from 71,936 principal investigators and observed diminishing returns in the annual weighted Relative Citation Ratio (RCR) beyond a GSI of 21, which is the equivalent of 3 R01 grants held by a single principal investigator (Figure 1).[5] While the attempt to optimize equity and productivity is laudable, the fairness of this data-driven cap depended squarely on the validity of the metrics and the quality of the data. We had several concerns.

The GSI is a metric that assigns points to grant types depending on their complexity and size.[6] The metric was designed by assigning seven points to the most common grant, the R01, and giving more points to larger grants and fewer points to the smaller ones.[5 6] This assignment of points lacked calibration: it is unclear how the GSI correlates with other metrics of grant support and whether more or lesser points need to be assigned to certain grant mechanisms.[2] The GSI should be considered an arbitrary point system until further analyses have demonstrated its validity.

The *annual weighted* RCR quantifies the output of a researcher's portfolio. This metric is based on the RCR, which indicates the influence of a single article, calculated as the ratio of the article citation rate and the expected citation rate for articles in its field.[7] The RCR is normalized against a large set of NIH publications so that a value of 1 indicates that the citation rate of an article is on par with the mean of articles in its field. The *weighted* RCR is the <u>sum</u> of RCRs for all articles published by a researcher in a certain period, and the *annual* weighted RCR is this sum divided by the number of publication years. The annual weighted RCR has shortcomings that may have introduced errors in setting the GSI threshold.

The calculation of the annual weighted RCR of researchers' portfolios assumes that the RCR of a scientific article is constant over time, but this is not the case:[8] the RCR decreases when an article is past its heyday. The citation rate of an article (the numerator in the RCR equation) decreases over time when its number of citations 'stabilizes' and the number of years keeps increasing. The expected citation rate (the denominator in the RCR equation) is a normalized *2-year* citation rate of the journals in which the 'field' articles were published. Using a 2-year citation rate assumes that the citation rate of an article is constant over time and always as high as in the first two years, which is unrealistic for older articles. A decreasing numerator and overestimated denominator result in reduced RCRs for most older articles.

When the RCRs of scientific articles decrease over time, the annual weighted RCR is not a suitable indicator for the *current* or *recent* scientific influence of mid- and late-career researchers, as their older work pulls down their portfolio RCRs. The reduction of their annual weighted RCR might, at least in part, explain the diminishing returns that were observed in the NIH data as researchers with higher GSI scores had substantially longer histories of grant funding (median of 19 years among GSI>21 as compared to 3 years among GSI≤7).[5] A metric that better reflects

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3106v1 | CC BY 4.0 Open Access | rec: 25 Jul 2017, publ: 25 Jul 2017

current or recent scientific influence of researchers' portfolios might have shown less diminishing returns, would have led to a higher cap and allowed more grants per researcher.

The legitimacy of the threshold is not only determined by the validity of the metrics, but also by the quality and validity of the data that were used to obtain the GSIs and annual weighted RCRs. The NIH clarified in detail how they calculated the GSI for each year, but not how they calculated nor which data they used to obtain the annual weighted RCRs.[5] Figure 1 shows that many scientists had annual weighted RCRs below 1, which means that their *total* annual output was not even the equivalent of one 'average' article with an RCR of 1. This low number raises questions about the data that were used to quantify the output in researchers' portfolios.

The NIH explained that they only considered publications in which grant funding was acknowledged, and they assigned publications only to the PI(s), not to co-authors,[5] but it is unknown how they determined the number of publication years to calculate *annual* weighted RCRs. They may have used the number of years between the first and last publication, the number of years with grant funding, or even the number of years for which publication data were available (1996-2014) because the exact number of publication years for each researcher could not be determined. The latter would overestimate the number of years for many researchers and could explain the low annual weighted RCRs.

An alternative explanation is that the low annual weighted RCRs were observed because the NIH used funding and publication data from the same period. Using data from the same period artificially increases the number of grants with no or little output because recently-awarded grants had not had the time to generate a competitive number of publications that were cited frequently enough to yield above average RCR ratings. The median lag time between the start of a grant and the first publication is three years, and many grants continue to deliver publications after a grant is completed.[9] When recent grants are removed from the NIH analysis, the size of the lowest GSI category, which set the expected output for the rest, will be strongly reduced as these researchers had a median of 3 years of funding.[5] Removing recent grants may reduce the expected returns from research funding for researchers with more grants and change the observation of diminishing returns.

Restricting the amount of support would have meant that scientifically outstanding grants, based on evaluation by study sections, would not be funded when they caused researchers to exceed their maximum allowable grant support. Overruling the award of competitive grants to successful researchers is a decision that should not be taken lightly. A data-driven threshold that affects the careers of scientists should be justified by nothing less than a rigorous analysis using valid metrics in high-quality data. As a science-based organization, the NIH values rigor, transparency, and high-quality data. These values must be upheld when developing metrics and policies that impact the scientific enterprise and the careers of scientists. The analytic justification of the GSI threshold did not meet that standard.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3106v1 | CC BY 4.0 Open Access | rec: 25 Jul 2017, publ: 25 Jul 2017

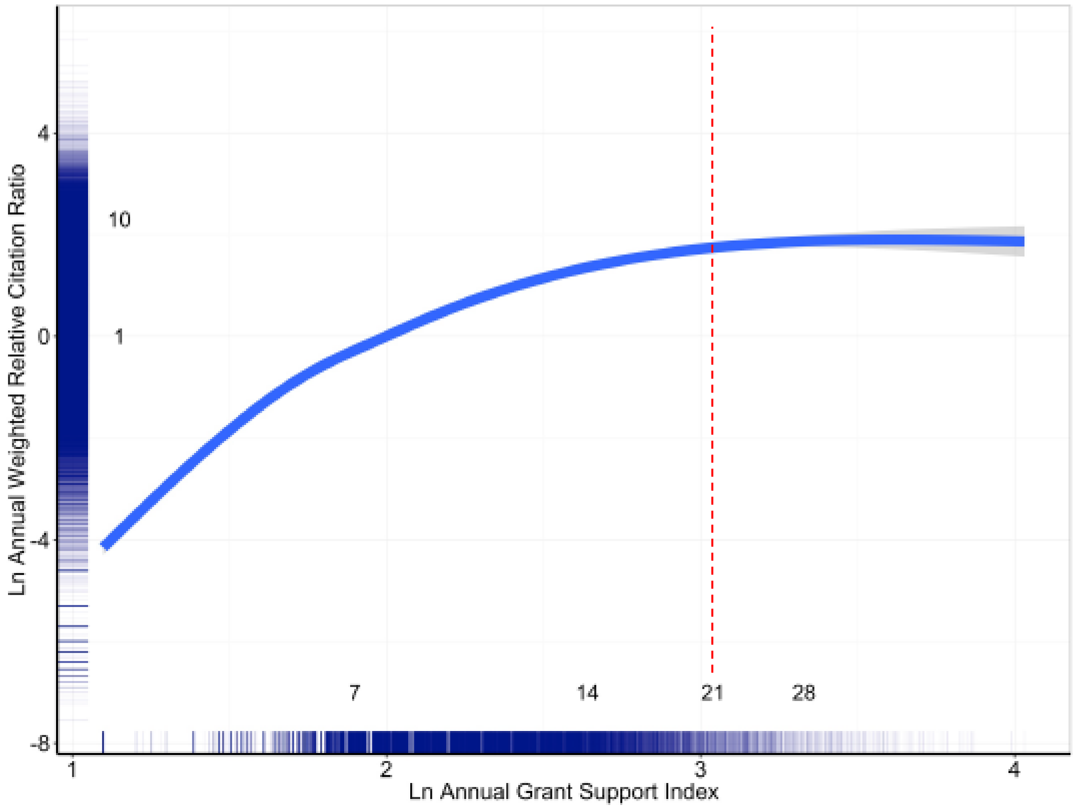
References

1. Collins FS. New NIH Approach to Grant Funding Aimed at Optimizing Stewardship of Taxpayer Dollars 2017 [Available from: https://www.nih.gov/about-nih/who-we-are/nih-director/statements/new-nih-approach-grant-funding-aimed-optimizing-stewardship-taxpayer-dollars accessed July 12 2017].
2. Lauer M. Implementing Limits on Grant Support to Strengthen the Biomedical Research Workforce 2017 [Available from: https://nexus.od.nih.gov/all/2017/05/02/nih-grant-support-index/ accessed July 12 2017].
3. Kaiser J. NIH scales back plan to curb support for big labs after hearing concerns. *Science* 2017(6338) doi: 10.1126/science.aan6901.
4. Lauer M. NIH's Next Generation Researchers Initiative 2017 [Available from: https://nexus.od.nih.gov/all/2017/06/16/nih-next-generation-researchers-initiative/ accessed July 12 2017].
5. Lauer M, Roychowdhury D, Patel K, et al. Marginal returns and levels of research grant support among scientists supported by the National Institutes of Health. *BioRxiv* 2017 [published Online First: May 26, 2017].
6. Lauer M. Research Commitment Index: A New Tool for Describing Grant Support 2017 [Available from: https://nexus.od.nih.gov/all/2017/01/26/research-commitment-index-a-new-tool-for-describing-grant-support/ accessed July 12 2017].
7. 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(9):e1002541. doi: 10.1371/journal.pbio.1002541.
8. Janssens ACJW, Goodman M, Powell KR, et al. A critical evaluation of the algorithm behind the Relative Citation Ratio (RCR). *PLoS Biology* 2017:In press.
9. Boyack KW, Jordan P. Metrics associated with NIH funding: a high-level view. *J Am Med Inform Assoc* 2011;18(4):423-31. doi: 10.1136/amiajnl-2011-000213.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3106v1 | CC BY 4.0 Open Access | rec: 25 Jul 2017, publ: 25 Jul 2017

Figure 1. **Grant Support Index and annual weighted Relative Citation Ratio for 71,936 principal investigators on NIH grants**

The graph was originally published in [5] (under CC-BY 4.0 International License). The numbers inside the axes represent the non-transformed values of the annual weighted RCR and GSI. A dashed red line was added to indicate the proposed GSI threshold.



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3106v1 | CC BY 4.0 Open Access | rec: 25 Jul 2017, publ: 25 Jul 2017