Journal of Informetrics 20 (2026) 101847



Contents lists available at ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi



The illusion of scientific talent identification through publication counts: a critical commentary on Haunschild and Bornmann

ARTICLE INFO

Keywords:
Bibliometrics
Research evaluation
Early-career researchers
Journal impact factor
Citation analysis
Responsible research assessment
Publication metrics

ABSTRACT

This letter responds to the article by Haunschild and Bornmann, published in the *Journal of Informetrics*, which proposes a bibliometric framework for identifying promising early-career scientists based on three specific indicators: the number of publications in top-quartile journals (Q1), total publication output (O), and corresponding authorship (C). Using the combined OxQ1 indicator, the authors compile a dataset of 46,200 individuals classified as "potentially talented," intended to support informed peer review and hiring decisions. This commentary directly examines the methodological assumptions underlying the authors' indicator selection and their validation strategy. The use of Q1—a journal-level quartile metric—as a proxy for individual research quality is problematic, particularly given the authors' own acknowledgment of the Journal Impact Factor's limitations. More critically, the authors' own validation analysis yields type I and type II error rates of approximately 28–30%, raising serious questions about the OxQ1 combination's reliability for the talent identification purposes proposed. The reliance on total publication output (O) as a core indicator risks conflating productivity with intellectual contribution, while the public release of a dataset labeling 46,200 researchers as "potentially talented" carries methodological and ethical implications that this letter addresses directly. The argument is not that bibliometrics lacks value, nor that Q1 journals lack quality, but that talent is a capacity which cannot be read off counts of output and journal placement; the letter explains why and points to indicators better suited to assessing individuals.

1. Haunschild and Bornmann's indicator selection: conceptual and empirical limitations

The article by Robin Haunschild and Lutz Bornmann, titled Identification of Potential Young Talented Individuals in the Natural and Life Sciences: A Bibliometric Approach, published in the Journal of Informetrics (Haunschild & Bornmann, 2023), proposes a bibliometric method for identifying promising early-career scientists based primarily on early publication performance. The authors rely on three indicators: the number of papers published in high-impact journals (Q1), the total number of papers produced (O), and the number of publications in which the researcher appears as corresponding author (C). Using these indicators, they compile a dataset of 46,200 individuals identified as "potentially talented," which they suggest could support informed peer-review processes and potentially assist hiring or evaluation decisions. Although the motivation to identify emerging scientific talent is understandable, the methodological assumptions underlying this proposal raise significant conceptual, empirical, and ethical concerns. In particular, the article reinforces a model of scientific evaluation that equates talent with publication productivity and journal prestige—two indicators whose limitations have been widely documented in the literature on research assessment. By reducing the complex phenomenon of scientific creativity and intellectual contribution to a small set of easily measurable outputs, such approaches risk oversimplifying the dynamics of scientific development and academic career trajectories. The analysis below examines each of these choices in turn, focusing specifically on the conceptual validity of the Q1 indicator, the predictive reliability of the OxQ1 combination as revealed by the authors' own validation data, and the implications of treating total publication output (O) as a primary measure of individual scientific talent. A prior question, however, concerns what is being measured at all. Talent denotes a capacity—an intellectual ability, a potential to create something that did not previously exist—and a potential of this kind is not the same thing as a record of outputs already produced. The number of papers in Q1 journals, the total count of publications, and the tally of

https://doi.org/10.1016/j.joi.2026.101847
Received 28 April 2026; Received in revised form 8 June 2026; Accepted 29 June 2026
1751-1577/© 2026 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

corresponding authorships are, at best, surrogate indicators of an individual's publication production and publication strategy; they do not measure the capacity that the word talent properly names. The distinction matters because talent, so understood, is a property of individuals rather than of large aggregates of scientists, and an indicator set assembled to sort tens of thousands of researchers is poorly suited to capturing it.

A first concern involves the use of publication in high-impact journals as a signal of scientific quality or individual potential. Journal impact factors are aggregate, journal-level metrics that do not measure the quality of individual articles or the specific contributions of researchers. Long-standing critiques of impact factors emphasize that they are highly skewed, field-dependent, and vulnerable to manipulation (Garfield, 2006). Using them as a proxy for individual talent therefore risks committing a well-known fallacy: inferring the value of an individual contribution from the average citation performance of the journal in which it appears. In practice, highly cited journals may publish articles that receive few citations, while influential articles may appear in journals with more modest impact metrics. Consequently, equating publication venue with research quality introduces substantial noise into evaluation processes. This concern applies directly to Haunschild and Bornmann's Q1 indicator, which assigns a binary value—within or outside the top quartile—based on a journal's field-normalized citation impact. Notably, the authors themselves cite the San Francisco Declaration on Research Assessment (DORA, 2012) and acknowledge that "the practice of basing selection decisions on the JIF has been heavily criticized." The inconsistency between acknowledging this limitation and then employing a derivative journal-tier indicator as the primary talent selection criterion warrants explicit justification, particularly when the intended downstream applications include hiring and fellowship decisions. It is useful to separate two levels at which this indicator operates. At the level of journals, the top quartile does identify, in most fields, the leading periodicals: the great majority of Q1 titles are of high standing, and the Journal Impact Factor, whatever its faults, remains one of the more serviceable journal-level indicators (Larivière & Sugimoto, 2019). The difficulty is that this validity does not carry over to the level at which Haunschild and Bornmann apply it. The share of an author's papers in Q1 journals describes a publication strategy and reflects the visibility a venue can offer; it is not an impact index, and it should not be weighted as though it were a relative citation indicator (Waltman & Traag, 2021). Scientometric indicators do not measure quality directly in any case; what they can assess is the impact of the information published, and that impact is revealed only after publication, through citations. A more defensible individual-level alternative, available for the validation cohort, is the number—or share—of an author's papers among the top 1%, 5%, or 10% most cited in their field.

Haunschild and Bornmann justify their avoidance of citation-based indicators on the grounds that such metrics "usually need too long to accrue" for recent researchers. This practical constraint is acknowledged in the scientometric literature. However, the authors' own validation findings complicate this rationale in an important way. As shown in Table 3 of the article, the Q1 indicator outperforms both O and C as a predictor of future citation impact—precisely because Q1 correlates with subsequent citation performance. This means Q1 functions in the framework primarily as a citation impact surrogate, which is only useful insofar as it anticipates what direct citation metrics would eventually reveal. For the validation cohort (first publications 1999–2003), citation data were available over a sufficiently long window; yet no direct comparison between Q1 and citation-based alternatives is reported. The absence of such a comparison limits the strength of the methodological validation and leaves unanswered the key question of whether Q1 adds predictive value beyond what could be captured by citation data directly. There is a further tension in setting citations aside. Q1 status is itself defined by field-normalized citation impact, so the framework does not so much avoid citations as admit them indirectly, in the aggregated form of the journal rather than the article. The stated reason for leaving citations out, that too little time has passed for them to accrue, also fits the cohort less evenly than it first appears. It is reasonable for someone in the first years after a recent debut, but the same top-1% threshold admits researchers whose follow-up record already runs into the hundreds of papers, up to the maximum of 1070 noted above; an output on that scale does not describe a career still too early for citation data to exist. Where a citation history has had time to form, the record those individuals have already earned is hard to set aside in favour of a journal-level proxy.

A further and more fundamental concern relates to the reliability of the proposed method as documented by the authors' own results. The article reports type I errors of 28.7% (individuals selected as talented who turn out to have low subsequent citation impact) and type II errors of 30.1% (control-group scientists who turn out to have high subsequent citation impact) when the OxQ1 indicator is applied to the validation cohort. These figures mean that roughly one in three researchers is misclassified relative to their subsequent citation impact. These error rates are disclosed honestly in the paper but are not prominently discussed in relation to the dataset's intended applications. The authors suggest that the dataset "can be used as a starting point to search for more information on selected talented individuals in certain disciplines" and propose it as a pre-selection resource for institutions. When a classification framework producing misclassification rates of this magnitude is made publicly available for use in competitive academic hiring, the risk of consequential false positives and false negatives warrants far more explicit acknowledgment than the article provides. These error rates also point to a more basic mismatch between the method and its stated aim. A sound assessment presupposes that the purpose of the exercise has been settled, that indicators suited to that purpose have been chosen, and that the relevant data are at hand; only then can the result be trusted, and the OxQ1 procedure is not well matched to the aim of detecting talented young scientists. Sorting individuals into the two classes of talented and not-talented does not meet this standard, because the influence scientists exert on their field is not dichotomous but distributed—approximately continuously, whatever the precise shape of the distribution. Forcing that continuum into two groups at a top-1% threshold is a coarse approximation, and clustering researchers according to the values of the indicators would describe the underlying gradation far more faithfully than a single cut-off.

Moreover, empirical evidence suggests that the prestige associated with high-impact journals does not necessarily correlate with research reliability or integrity. Arturo Casadevall and Ferric C. Fang (2014), in an analysis published in *mBio*, found that articles retracted due to data falsification or fabrication are disproportionately concentrated in high-impact journals. This observation highlights a striking paradox: the same journals often used as indicators of excellence in evaluation systems may also be environments

where intense competition and strong career incentives increase the risk of problematic research practices. By selecting Q1 journal placement as a primary indicator, Haunschild and Bornmann's framework implicitly rewards publication in the high-impact venues where, as Casadevall and Fang demonstrate, competitive pressures and strong career incentives are most concentrated. This is a direct consequence of using journal prestige as the principal quality signal for talent identification rather than an incidental limitation.

Similarly, Grimes et al. (2018) argue that the contemporary scientific reward system—particularly the strong emphasis on publishing—can generate incentives for careless or even fraudulent research practices. Leading journals frequently prioritize novel and positive findings over studies reporting null results, a tendency that may contribute to the broader reproducibility crisis and ultimately undermine public trust in science. This concern is directly relevant to Haunschild and Bornmann's inclusion of total output (O) as a core indicator. Their Table 2 shows that the "potentially talented" group identified via O alone has a median of 42 subsequent publications, compared to 14 in the control group. To the extent that high output partly reflects optimization for publication frequency rather than depth of contribution, this validation criterion may confirm productivity rather than genuine scientific quality.

Haunschild and Bornmann do not discuss any upper bound on meaningful individual publication rates implied by their O indicator. In the validation cohort, the maximum publication count in the ten-year follow-up period reaches 1070 papers (Table 2)—a figure that corresponds to roughly two publications per week sustained over a decade. The inclusion of such extreme output rates within the top-1% talent threshold raises a concrete methodological question: whether the O indicator, as operationalized, consistently measures scientific productivity in the conventional sense or whether it captures, in some portion of cases, authorship practices that inflate individual publication counts.

Beyond their conceptual limitations, metrics that equate talent with publication volume may contribute to broader systemic distortions within the scientific ecosystem. When academic evaluation systems reward quantity of output, researchers are incentivized to maximize publication counts rather than pursue riskier, potentially transformative research. This dynamic contributes to the accelerating expansion of the scientific literature, a phenomenon that has itself become the subject of critical investigation. In a large-scale study examining 1.8 billion citations across 90 million papers in 241 subject areas, Chu and Evans (2021) concluded that "a deluge of papers does not lead to turnover of central ideas in a field, but rather to ossification of canon." In other words, the rapid proliferation of publications does not necessarily foster intellectual renewal; instead, it may entrench established paradigms by overwhelming the scholarly attention required to identify genuinely influential contributions. This dynamic is directly relevant to the role of O in Haunschild and Bornmann's framework. By identifying researchers who rank among the top 1% in total early-career output, the OxQ1 combination tends to select those who have most successfully pursued high-volume publication strategies. This may not consistently identify those whose work makes the most distinctive intellectual contributions—which is precisely the goal the authors state.

2. Authorship indicators, validation reliability, and the ethics of publicly classifying researchers

The strong emphasis placed on publication counts in many research evaluation systems may also contribute to the emergence of hyper-prolific authorship. In a widely cited analysis, Ioannidis et al. (2018) defined hyper-prolific scientists as those publishing approximately one paper every four days—and identified >9000 authors worldwide who met this threshold during the period analyzed. Such extraordinary publication rates raise important questions about authorship practices, the distribution of intellectual contribution within large research teams, and the incentives embedded in contemporary academic reward structures. In the context of

Table 2
Statistical values of the number of papers for each group and indicator combination since ten years following the scientists' first publication until 2018.

Group	Indicator combination	Minimum	1st quartile	Median	Mean	3rd quartile	Maximum
Potentially talented individuals	O	1.00	21.00	42.00	59.80	77.00	1070.00
	Q1	1.00	15.00	34.00	51.57	65.00	1070.00
	C	1.00	16.00	34.00	48.08	63.00	765.00
	OxQ1	1.00	28.00	55.00	77.55	97.00	1070.00
	OxC	1.00	26.00	49.00	65.09	85.00	765.00
	Q1xC	1.00	27.00	52.00	68.90	90.00	765.00
	OxQ1xC	1.00	35.00	62.00	80.24	104.00	765.00
Control group	O	1.00	6.00	14.00	23.58	30.00	752.00
	Q1	1.00	6.00	16.00	27.69	36.00	752.00
	C	1.00	6.00	15.00	26.07	33.00	760.00
	OxQ1	1.00	8.00	19.00	29.70	38.00	599.00
	OxC	1.00	8.00	18.00	27.99	36.00	529.00
	Q1xC	1.00	9.00	22.00	33.25	43.00	532.00
	OxQ1xC	1.00	10.75	24.00	34.18	44.00	499.00
Differences between potentially talented individuals and control group	O	0.00	15.00	28.00	36.22	47.00	318.00
	Q1	0.00	9.00	18.00	23.88	29.00	318.00
	C	0.00	10.00	19.00	22.02	30.00	5.00
	OxQ1	0.00	20.00	36.00	47.86	59.00	471.00
	OxC	0.00	18.00	31.00	37.10	49.00	236.00
	Q1xC	0.00	17.00	30.00	34.64	45.00	233.00
	OxQ1xC	0.00	23.00	38.00	44.89	59.00	376.00

Table 3
Statistical values of the Hazen percentiles of papers for each group and indicator combination since ten years after the scientists' first publication until 2018.

Group	Indicator combination	1st quartile	Median	Mean	3rd quartile
Potentially talented individuals	O	43.41	57.59	56.45	70.80
	Q1	55.96	68.32	66.45	79.20
	C	40.39	55.29	53.87	68.89
	OxQ1	56.07	67.84	66.27	78.46
	OxC	43.67	57.53	56.04	70.17
	Q1xC	55.56	67.15	65.31	77.15
	OxQ1xC	55.87	67.13	65.47	77.06
Control group	O	36.69	53.10	52.20	68.29
	Q1	44.58	59.27	57.94	72.78
	C	35.18	51.65	50.81	66.84
	OxQ1	44.14	57.96	56.64	70.54
	OxC	37.37	52.66	51.95	67.19
	Q1xC	43.78	57.92	56.59	70.60
	OxQ1xC	44.38	58.04	56.61	70.47
Differences between potentially talented individuals and control group	O	6.72	4.49	4.24	2.51
	Q1	11.37	9.05	8.51	6.41
	C	5.21	3.65	3.06	2.05
	OxQ1	11.93	9.88	9.64	7.92
	OxC	6.31	4.87	4.09	2.98
	Q1xC	11.78	9.23	8.73	6.55
	OxQ1xC	11.49	9.08	8.86	6.59

Haunschild and Bornmann's framework, this is not a theoretical concern. Authors publishing at rates approaching one paper every few days will rank among the top 1% by total output (O) over a five-year window almost by construction, and will therefore appear in the dataset of 46,200 "potentially talented" individuals—regardless of whether their publication rates reflect genuine scientific productivity or systematic authorship inflation. The O indicator, as operationalized, provides no mechanism for distinguishing between these cases.

Concerns about extreme publication productivity have been raised by several scholars. For example, Steven L. Salzberg has argued that the unusually high output of the microbiologist Didier Raoult was partly explained by the practice of attaching his name to nearly every paper produced by his institute—a practice Salzberg characterized as "grossly unethical" (Pacheco-Torgal, 2021). This concern applies specifically to the corresponding authorship indicator (C) employed by Haunschild and Bornmann. When C is operationalized as the count of papers in which a researcher appears as corresponding author, it may capture institutional authority or laboratory leadership rather than direct intellectual contribution. Authors at the head of large research groups may accrue high C values precisely through the kind of systematic attachment of names to group output that Salzberg characterizes as problematic. The C indicator as used in this study does not distinguish between corresponding authorship that reflects intellectual leadership and corresponding authorship that reflects institutional hierarchy. A related point concerns how credit is allocated across the by-line. To approximate the share of the work actually performed by a young researcher, the credit for a paper has to be divided among its coauthors rather than assigned in full to each name on it; Haunschild and Bornmann undertake no such allocation, although it bears directly on the evaluation of early-career scientists. Junior researchers are, as a rule, not in a position to appear as first or corresponding author even when they have carried out a substantial part of the work: their place in the list of authors is governed largely by local conventions, and a corresponding-author role is seldom available to someone whose position in a laboratory is not yet secure. Several established methods assign fractional credit by rank in the by-line, harmonic counting among them (Hagen, 2008), and at a minimum any individual-level evaluation should take the number of authors and their order into account.

Evidence from the research evaluation industry also underscores the tension between publication volume and meaningful scholarly contribution. A report by Clarivate noted that authors publishing at rates of two or three papers per week challenge conventional understandings of authorship and credit allocation within the scientific community (Halevi, 2022). Evaluation systems that prioritize numerical output risk legitimizing practices in which authorship becomes strategically distributed across large collaborative networks, thereby complicating the attribution of intellectual responsibility and weakening traditional norms of scholarly accountability.

Another important concern involves the structural inequalities that bibliometric indicators may reinforce. Access to prestigious publication venues is often mediated by institutional resources, mentorship networks, and geographic location. Researchers working in well-funded laboratories or internationally recognized universities typically benefit from stronger collaborative infrastructures, greater visibility within global scientific networks, and better support for publication activities. Consequently, publication-based indicators may capture institutional advantage as much as individual talent. This concern is particularly salient in Haunschild and Bornmann's framework given their use of Q1 as a central indicator. Access to top-quartile publication venues is systematically mediated by institutional resources, English-language proficiency, international collaboration networks, and research funding levels. Researchers working in low-income countries or at institutions without strong international visibility are structurally disadvantaged by a Q1-based selection criterion, independent of the quality of their scientific work. The dataset of 46,200 individuals, derived from Scopus and weighted toward Q1 publications, is therefore likely to overrepresent researchers from well-resourced institutional contexts in high-income countries. This limitation is not directly addressed in the authors' discussion of the framework's intended

applications. A clarification is in order here. The aim of scientometric impact measurement is to register the impact that scientific work has had, without conditioning that measurement on the funding, language, age, or other circumstances of its authors; it would be a mistake to ask citation or impact indicators to be corrected for financial conditions, except perhaps in dedicated sociological studies. The argument advanced here is narrower and does not run against that principle. It is that once the stated object shifts from measuring realized impact to inferring individual talent, the structural mediation of access to Q1 venues enters as a confound between the indicator and the capacity it is meant to signal, so that what is recorded as Q1 productivity may in part register institutional advantage rather than the potential of the person.

At the same time, it is important to recognize that contemporary academic reward systems increasingly rely on mechanisms of recognition associated with scholarly influence rather than simple productivity counts. Lauren M. McIntyre, Editor-in-Chief of G3: Genes|Genomes|Genetics, has emphasized that citations constitute one of the primary mechanisms through which the global scientific community acknowledges intellectual contributions (McIntyre, 2024). Citations represent a form of epistemic recognition, indicating that a given piece of research has informed or shaped subsequent work. This perspective stands in contrast to Haunschild and Bornmann's deliberate choice to exclude citation-based indicators from their framework. The authors' own validation results—showing that Q1 predicts future citation impact better than O and C precisely because of its correlation with citation performance—suggest that citation-based indicators, where available, provide more direct evidence of the intellectual influence the framework aims to identify.

Consistent with this perspective, a recent study reported in *Nature* analyzing the careers of >40,000 researchers across eight scientific disciplines found that securing a research or faculty position at a top-tier university in a developed country is strongly associated with publishing at least one article among the top 5% most-cited papers during postdoctoral training (Ryan, 2025). This finding illustrates how citation-based recognition—not simply the number of publications—plays a significant gatekeeping role in academic career trajectories.

The public release of a dataset labeling 46,200 individuals as "potentially talented" requires careful consideration precisely in light of the error rates the authors report. The OxQ1 combination, as the authors' own Table 4 demonstrates, misclassifies approximately 30% of researchers relative to their subsequent citation impact. Haunschild and Bornmann explicitly note that their dataset "cannot and should not replace established institutional selection processes" and that it provides only "hints of possibly talented individuals." However, once a freely downloadable list of individuals labeled as "potentially talented" enters institutional practice, the careful caveats of the academic paper may not reliably accompany its use. Institutions or funding bodies with limited bibliometric expertise may treat the dataset as a validated ranking rather than a probabilistic pre-selection with a 28–30% error rate. The authors do not discuss this translation risk, which is arguably the most consequential ethical dimension of the work.

Beyond how it is used in practice, the dataset's very existence raises epistemic and ethical questions. Bibliometric indicators, particularly those based primarily on publication counts or early-career productivity, are inherently noisy and context-dependent. Numerous influential scientists produced relatively modest publication records during the early stages of their careers, while others with strong early productivity did not necessarily go on to generate transformative contributions. Classifying individuals as "potentially talented" on the basis of a limited set of productivity-oriented indicators therefore risks generating both false positives and false negatives. Haunschild and Bornmann quantify these errors explicitly: 28.7% type I and 30.1% type II. These figures should be treated not as a minor caveat but as a central feature of the framework's performance that is communicated prominently to any institution downloading the dataset for selection purposes.

None of this is to suggest that citations offer a flawless remedy. Citation counts carry their own well-known defects—disciplinary differences in citing behaviour, self-citation, the lag before impact becomes visible, and citations that are perfunctory or even critical rather than endorsing—so they cannot simply be substituted for the indicators criticized here and treated as decisive on their own. The reliability of any of these measures is, moreover, lowest precisely where Haunschild and Bornmann apply them: the smaller the body of work to be judged, as it necessarily is for early-career researchers, the harder a defensible scientometric judgement becomes. For such individuals, scientometric indicators are best read alongside peer review rather than in place of it, and altmetric indicators, which register attention and use more quickly than citations accrue, have been proposed as a complement suited to early-career assessment (Bornmann, 2014). The difficulty eases considerably for larger units of analysis, such as the lifetime output of established scientists or the aggregate record of teams, institutions, and countries, where the larger number of items makes the indicators more stable.

These concerns do not imply that bibliometric analysis has no legitimate role in the study of scientific activity. On the contrary, bibliometrics has proven highly valuable for examining collaboration networks, disciplinary evolution, and the diffusion of knowledge. However, the gap between bibliometrics' descriptive and analytical value on the one hand, and its use as an individual-level talent classification tool on the other, is precisely the gap that the framework proposed by Haunschild and Bornmann attempts to bridge—and the gap that the concerns raised here suggest has not yet been fully closed. Rather than reinforcing an evaluation model centered primarily on publication counts and journal prestige, future research in the field of scientometrics could contribute more constructively by developing approaches that integrate qualitative assessment, field-normalized citation indicators, and contextual information about research environments and career trajectories. It would be wrong, though, to read the foregoing as confining scientometrics to such descriptive and analytical uses. Assessing the use, merit, and impact of research—of individuals as much as of teams, disciplines, and countries—is one of its central tasks, and applied scientometrics underwrites a great deal of practical science policy: grant decisions, academic appointments, and scholarly distinctions are rarely made today without recourse to indicators of this kind. The literature already offers tools fit for individual-level assessment, among them the h-, g-, and π-indices, the I3 indicator, percentile-based measures such as the share of papers in the top 1% or 10% of a field, and indicators such as PR(2,10) or PR(2,50) (Bornmann et al., 2011; Ioannidis et al., 2019). What the present case illustrates is not that individuals should not be evaluated, but that evaluation must begin from a clearly defined aim, with methods, data, and indicators chosen to fit it (Moravcsik, 1988), and

should be carried out by those with scientometric expertise—conditions the OxQ1 proposal does not adequately meet.

3. Conclusion

The proposal by Haunschild and Bornmann to identify promising early-career researchers through the OxQ1 bibliometric indicator is motivated by a legitimate practical concern: competition for positions in academic science is intense, candidate pools are large, and efficient pre-selection tools have genuine administrative value. The authors are also commendably transparent about some of the limitations of their approach, including the imperfect author disambiguation in Scopus and the possibility that talented individuals may be missed. Nevertheless, the analysis presented here identifies three specific concerns arising directly from the methodological choices made in the article.

First, the Q1 indicator—while operationally convenient—inherits the well-documented conceptual limitations of the Journal Impact Factor. The authors themselves cite both the DORA declaration and Garfield's own critique of the JIF; employing a derivative journal-tier metric as the primary quality signal in a talent identification framework intended to inform hiring decisions sits in tension with these acknowledgments. The article would benefit from a more direct engagement with this tension and a clearer justification for why Q1 is preferred over citation-based alternatives for the validation cohort, where citation data were available.

Second, the OxQ1 combination produces type I and type II error rates of approximately 28–30% in the authors' own validation analysis. This level of misclassification—roughly one in three researchers incorrectly classified—should be foregrounded, not footnoted, in the discussion of the dataset's intended applications. Institutions downloading the dataset should be clearly informed that the framework performs at this level of accuracy, so that the "hints" the authors intend are not treated as validated talent rankings.

Third, the reliance on total output (O) and corresponding authorship (C) as core indicators risks systematically identifying patterns of hyper-prolific publication and institutional leadership rather than individual intellectual distinction. Researchers with the highest early-career publication counts are not, by definition, those making the most original contributions; nor are those accumulating the most corresponding authorships necessarily those exercising the most direct intellectual responsibility. These are not merely theoretical objections—they reflect concrete limitations of the operationalization choices made in the study.

Addressing these limitations would substantially strengthen future iterations of the proposed framework and would help ensure that bibliometric talent identification tools contribute to fairer and more accurate research assessment rather than reinforcing the productivity-oriented incentive structures they are meant to complement.

References

Bornmann, L., Mutz, R., Hug, S. E., & Daniel, H. D. (2011). A multilevel meta-analysis of studies reporting correlations between the h index and 37 different h index variants. *Journal of Informetrics, 5*(3), 346–359.

Bornmann, L. (2014). Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of altmetrics. *Journal of Informetrics, 8*(4), 895–903.

Casadevall, A., & Fang, F. C. (2014). Causes for the persistence of impact factor mania. *mBio, 5*(2), Article e00064-14.

Chu, J. S., & Evans, J. A. (2021). Slowed canonical progress in large fields of science. In *, 118. Proceedings of the National Academy of Sciences*, Article e2021636118.

DORA. (2012). San Francisco declaration on research assessment.

Garfield, E. (2006). The history and meaning of the journal impact factor. *JAMA, 295*(1), 90.

Grimes, D. R., Bauch, C. T., & Ioannidis, J. P. A (2018). Modelling science trustworthiness under publish or perish pressure. *Royal Society Open Science, 5*(1), Article 171511.

Hagen, N. T. (2008). Harmonic allocation of authorship credit: Source-level correction of bibliometric bias assures accurate publication and citation analysis. *PLoS ONE, 3*(12), Article e4021.

Halevi, G. (2022). Why misconduct could keep scientists from earning highly Cited Researcher designations, and how our database plays a part. *Retraction Watch*.

Haunschild, R., & Bornmann, L. (2023). Identification of potential young talented individuals in the natural and life sciences: A bibliometric approach. *Journal of Informetrics, 17*(3), Article 101394.

Ioannidis, J. P., Klavans, R., & Boyack, K. W. (2018). Thousands of scientists publish a paper every five days. *Nature, 561*(7722), 167–169.

Ioannidis, J. P. A., Baas, J., Klavans, R., & Boyack, K. W. (2019). A standardized citation metrics author database annotated for scientific field. *PLoS Biology, 17*(8), Article e3000384.

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* (pp. 3–24). Cham: Springer.

McIntyre, L. M. (2024). Science is a team sport: Citations are how we recognize members of the team. *G3: Genes, Genomes, Genetics, 14*(7), Article jkae102.

Moravcsik, M. J. (1988). Some contextual problems of science indicators. In A. F. J. van Raan (Ed.), *Handbook of quantitative studies of science and technology*. Amsterdam: Elsevier/North-Holland.

Pacheco-Torgal, F. (2021) How many papers can a super-scientist author in a single week ? https://pacheco-torgal.blogspot.com/2021/03/how-many-papers-can-superscientist.html.

Ryan, J. (2025). More than 40% of postdocs leave academia, study reveals. *Nature*. https://doi.org/10.1038/d41586-025-00142-y

Waltman, L., & Traag, V. A. (2021). Use of the journal impact factor for assessing individual articles: Statistically flawed or not? *F1000Research, 9*, 366.

F. Pacheco-Torgal

C-TAC Research Centre, University of Minho, Campus de Azurem, 4800-058 Guimarães, Portugal

E-mail address: torgal@civil.uminho.pt.