# A Critical Analysis of the World's Top 2% Most Influential Scientists: Examining the Limitations and Biases of Highly Cited Researchers Lists

Akira J Abduh[1]

[1]Affiliation not available

January 22, 2023

## Abstract

This paper analytically analyzes the use and limitations of ranking systems for highly cited researchers compiled by Stanford University's ranking of the world's top 2% most influential scientists. This list is commonly used to identify influential and respected members of a particular field. However, it is important to critically evaluate the list and its methodology and no such analysis to this date. From a critical analysis of the September 2022 version of this world's top 2% of scientists list, this research finds that the database of the list is flawed, including inaccurately listing researchers as first publishing in the $19^{\text{th}}$ century and continuing to publish until 2022, listing authors with low publication numbers and career lengths, mixing news articles and editorials with research papers, listing institutes as authors rather than individuals, and listing authors with a high percentage of self-citations. The study suggests that the promotion and use of such "standardized" citation rankings should be discouraged.

## Introduction

Highly cited researchers are scientists or scholars who have published a significant number of papers that have been cited by other researchers in their field. These individuals are considered to be experts in their field and their research is often considered to be highly influential and important. The use of highly cited researchers is to identify influential and respected members of a particular field, and to help identify key research areas and trends within that field. Additionally, highly cited researchers is also referred by institutions and funding agencies to identify and support researchers who are making important contributions to their field, as well as by journals to identify potential reviewers and editorial board members.

One of the most foremost list is Highly Cited Researchers compiled by the Institute for Scientific Information (ISI) that identifies scientists and social scientists who have published a high number of papers that are among the top 1% most cited in their respective fields. The list is based on data from the Web of Science, a database that indexes and tracks scientific literature and conference proceedings. The Highly Cited Researchers list is

updated periodically, and it is considered to be a measure of the impact and influence of a researcher's work within their field. Scientists who are included on the list are considered to be among the leading researchers in their field, and their work is often considered to be highly influential and important.

As a competition, Stanford University has published a new ranking of the world's top 2% most influential scientists across all fields using Scopus citation data [1,2]. This database, created by Prof. John PA Ioannidis and his team at Stanford, is widely regarded as the largest collection of citation data and the most esteemed of its kind. This ranking aims to avoid misuse and misinterpretation of citation metrics as a measure of impact or excellence. To make it transparent and supposedly better than the ISI Highly Cited Researchers, the Stanford ranking uses "standardized citation metrics" to systematically rank the most-cited scientists in each field and provides standardized information on citations, h-index, co-authorship adjusted hm-index, and a composite indicator (c-score) to rank the top scientists in the world [1].

The top 2% or Stanford list was first published in 2019 to provide a ranking of scientists who have published a high number of papers that are among the top 2% most cited in their respective fields. The list is updated annually and the latest version, version 5, was published in September 2022 [1,2,3].

This Top 2% list is widely considered to be a prestigious list and is often used as a metric of a researcher's productivity and influence in their field. However, it is important to evaluate the list and its methodology critically. Some criticisms of this kind of highly cited list include that it relies heavily on the number of citations a researcher has received, which may not always be an accurate measure of the quality or impact of their work. Additionally, the list may be biased towards researchers from developed countries and institutions with strong research programs. Another critique is that the list may be biased towards quantity over quality of research, and having many publications does not necessarily mean that the researcher is doing high-quality research.

Despite its limitations, the Highly Cited Researchers and the Stanford Top 2% most influential scientists list is still widely used as a measure of a researcher's productivity and influence, and being named to the list is generally considered a significant achievement. Universities around the world proudly paraded their researchers when this list was announced every year. University news pompously reported that dozens of their researchers were included on Stanford University's list of the world's top 2% of scientists.

This top 2% list, is widely accepted without critical analysis. As there is no such analysis to this date, this study is the first to judiciously examine and critique the list by analyzing the outliers of the database. By studying the characteristics of this top 2% database, this study aims to gain insight into the factors that contribute to the database and identify potential misreporting and misuse. This paper goes all the way to suggest ways to improve it, to prevent further misuse. This is the first in the world known study that critically analyses the world's top 2% of scientists list.

# Methods

To investigate any potential outliers in the world's top 2% database of highly cited researchers, this study uses the "Updated science-wide author databases of standardized citation indicators" by Ioannidis, which can be found at https://elsevier.digitalcommonsdata.com/datasets/btchxktzyw. This database includes standardized information on citations, career-long and recent year impact for 22 scientific fields, and includes scientists who are ranked at the 2% or above percentile in the sub-field. The version 5 of this database is based on a Sept 1, 2022 snapshot from Scopus [3].

The database contains standardized information on citations, career-long and recent year impact for 22 scientific fields. The list includes scientists who are ranked at the 2% or above percentile in the sub-field. The version 5 of this database is based on a Sept 1, 2022 snapshot from Scopus.

The study analyzes the database in terms of the number of papers published, year of first publication, year of most recent publication, number of years of publishing, number of papers per year, number of single authored

papers, h-index as of end-2021, total cites 1996-2021, self-citation percentage, and no. cites/yr. The software MS Excel was used to compute the statistics (min, max, median) and identify statistical outliers.

# Results

## Statistic and outliers of the standardized citation indicators

The standardized citation list (version 5) contains 194,983 records of authors. The list has researchers ranked from no. 1 to 2,698,859 based on c-scores. Table 1 presents statistics of the authors in that database.

Table 1. Statistics of the databases of standardized citation indicators

| Parameters | Min | Max | Median | Description |
|---|---|---|---|---|
| np6021 | 2 | 3791 | 156 | # papers 1960-2021 |
| firstyr | 1834 | 2019 | 1986 | year of first publication |
| lastyr | 1960 | 2024 | 2022 | year of most recent publication |
| nyear | 2 | 187 | 35 | no. years of publishing (last-first yr) |
| nps (ns) | 0 | 2234 | 8 | number of single authored papers |
| np/yr | 0.02 | 220.3 | 4.7 | # papers/year |
| nps/yr | 0.00 | 217.7 | 0.2 | # single authored papers/yr |
| h21 (ns) | 3 | 284 | 35 | h-index as of end-2021 |
| nc9621 (ns) | 51 | 428620 | 4967 | total cites 1996-2021 |
| self% | 0.0% | 93.5% | 12.1% | self-citation percentage |
| ncites/yr | 2.7 | 35163.7 | 150.7 | no. cites/yr |

Potential outliers in this citation database may include authors with an excessive number of publications, a large number of citations compared to others in the same field, authors with a high number of self-citations, and authors that a small number of highly-cited papers have cited. Additionally, there may be authors that have been cited by a large number of low-quality publications. These outliers may be caused by factors such as the popularity of a particular research area, the quality of the research, the reputation of the authors or institutions, or even errors in the data. Based on Table 1, outliers in the data can be identified and grouped into several categories.

## Oldest and Youngest Researchers

One most prominent of outliers in the data is the oldest and most active researchers. In this study, the difference between the last and first year of publication is considered as the publishing age of a researcher. The median age in the database is 35 years, with a first publication in 1986 (Table 1). The database contains researchers who first published in the 19th century. Table 2 presents the top oldest and still active researchers (based on the year of most recent publication).

The oldest researcher in the Stanford list is William S. Marshall from St. Francis Xavier University in Canada, who was listed as first published in 1834 and continues to publish until 2021, which makes him 187 years of publishing history. However, this example highlights a flaw in the database, as it does not perform any rationality check to determine if an author can continuously publish for more than 80 years. The database lists 221 authors who have published for more than 80 years. In fact, William or Bill Marshall is a Senior Research Professor in Biology at St. Francis Xavier University. Bill had a Ph. D. from the University of British Columbia in 1977 and had Postdoctoral studies at the University of California at Berkeley. His first publication was in 1979 and is still active in research. His research is epithelial and renal physiology in lower vertebrates, especially teleosts.

The next oldest researcher listed is Lord Kelvin, who is listed to have published in 1849 until 2011, a total of 162 years with only 6 papers. Lord Kelvin's real name is William Thomson, or known as 1st Baron Kelvin, OM, GCVO, PC, PRS, FRSE (1824 –1907). In fact, he is a British scientist who died in 1907 and could not have published until 2011.

Francis Bonnet, an anesthesiologist from the Sorbonne University in France, was mentioned as having published his first work in 1866. This is incorrect, as Dr. Francis Bonnet, MD is a Professor at The Assistance Paris Hospitals (Ap-Hp) Seat. His research interests are Ambulatory Anaesthesia, Ambulatory Surgery and he is still alive.

Next in the list is University of Zurich's Franz E. Weber, a professor of dental medicine, was noted as having published his first paper in 1866. But this professor is still alive in Zurich with research in cranio-maxillofacial and oral surgery.

The top scientist John G.F. Cleland from the University of Glasgow, who has 1236 publications to his credit (since 1960) but was listed as having a first publication in 1867 and ranked at position 735. Professor Cleland is a professor of Cardiovascular & Metabolic Health still working in Glasgow.

These are just few samples of incorrect attribution list. In fact, this database lists 221 authors as having published for more than 80 years.

Table 2. Top oldest and most active researchers. authfull : author name; inst_name: institution name; firstyr: year of first publication; lastyr: last year of publication; np6021: # papers 1960-2021; rank (ns): rank based on composite score c, self-citations excluded; h21 (ns) h-index as of end-2021.

| authfull | inst_name | np6021 | firstyr | lastyr | rank (ns) | h21 (ns) |
|---|---|---|---|---|---|---|
| Marshall, William S. | St. Francis Xavier University | 114 | 1834 | 2021 | 48612 | 39 |
| Kelvin, Lord | University of Glasgow | 6 | 1849 | 2011 | 397622 | 8 |
| Bonnet, Francis | Sorbonne Universite | 418 | 1866 | 2022 | 92584 | 51 |
| Weber, Franz E. | Universität Zürich | 163 | 1866 | 2022 | 223953 | 33 |
| Cleland, J. G.F. | University of Glasgow | 1236 | 1867 | 2022 | 735 | 120 |

On the other hand, the Stanford list also includes some of the best and youngest scientists. Some of these "young researchers" with their first publications in 2019 and 2018 are listed in Table 3. Elisabeth Mahase from BMJ is an interesting case; she started publishing in 2019 and has already written 661 publications in 3 years while receiving 3135 citations. Elisabeth Mahase, in fact is not a scientist but a medical journalist working for BMJ. She is a news reporter at The BMJ (British Medical Journal). She reports on the news, not doing research. But this Stanford list mixed news article with peer review scientific articles. Mahase has topped most authors in scientific output, but she doesn't need to do any research.

Table 3 also shows various researchers who just published since 2018 but managed to be in the list. To make this list more interesting with weird entries, Table 4 shows some atypical authors who only had 2 years of publication record and published 5 to 12 papers. These few curious cases of authors in the top 2% list show the serious flaw of this Stanford list which needs fixing, it's broken.

Table 3. Top youngest researchers

| authfull | inst_name | np6021 | firstyr | lastyr |
|---|---|---|---|---|
| Mahase, Elisabeth | BMJ | 661 | 2019 | 2022 |
| Prata, Joana C. | Universidade de Aveiro | 39 | 2018 | 2022 |
| Davarpanah, Afshin | Aberystwyth University | 99 | 2018 | 2022 |
| Mikhaylov, Alexey | Financial University under the Government of the Russian Federation | 59 | 2018 | 2022 |
| Alola, Andrew Adewale | Vaasan Yliopisto | 93 | 2018 | 2022 |

Table 4. Researchers with only 2 years of publication career.

| authfull | inst_name | np6021 | firstyr | lastyr | rank (ns) | h21 (ns) |
|---|---|---|---|---|---|---|
| Tian, Jun | Digimarc Corporation | 6 | 2001 | 2003 | 213544 | 4 |
| Shi, Hang | Valence Technology | 7 | 1996 | 1998 | 311032 | 7 |
| Lee, Chuen Chien | University of California, Berkeley | 5 | 1989 | 1991 | 119202 | 5 |
| Patir, N. | Northwestern University | 12 | 1977 | 1979 | 244559 | 7 |

## Most Hyper Prolific and Most Unproductive Researchers

According to the number of publications and the number of papers produced annually, Table 5 includes some of the most hyper prolific authors. It is noteworthy that during the course of their careers, the top authors produced more than 3000 papers, which is amazing [4,5]. Actually, 930 of the list's authors have more than 1000 papers to their credit. Table 5 further breaks down the quantity of papers by the duration of the study (calculated by last minus first year). Elisabeth Mahase once more emerged as the most prolific author, producing nearly 220 papers annually.

Some of the authors (Table 5) are journalists, according to a more thorough investigation. For BMJ, Elisabeth Mahase and Abi Rimmer cover news. Even more alarmingly, Mahase and Rimmer outperformed most authors in terms of scientific output, since Scopus regards news articles as "peer review" publications.

Table 5 includes John P.A. Ioannidis, who has critically published about hyper prolific authors despite not being in the top 5. Since 1994, Ioannidis has published 45 papers per year on average. According to Scopus, Ioannidis published 52-80 papers per year from 2016 to 2021, or 1 paper every 5 days, "a figure that many would consider implausibly prolific" [4,5]. According to this ranking, Ioannidis is ranked 32 in the database, higher than the majority of Nobel laureates.

The most prolific authors who have written a paper entirely by themselves are further examined in this study. Table 6 shows the outcomes. It's incredible that these authors, who are all from clinical medicine, have written almost 2200 publications by themselves. The majority of these authors are journalists from BMJ and other news organizations, as was previously mentioned. The exception is Viroj Wiwanitkit from Dr. D. Y. Patil Vidyapeeth, a University in Pune who published, on average, 74 papers per year.

This outcome emphasizes even more the basic issue in so-called "databases of standardized citation indicators," which fail to distinguish between news pieces and articles written by journalists [3]. This so-called standardized citation indicator gives first-author and single-authored papers more weight. Scientists from MIT, Stanford, or any other scientist who had to put in a lot of effort in experimental research have a considerably lower c rank than writer Elisabeth Mahase, who can produce a one-page news piece every two days and is listed with a rank of #33363. According to the Stanford list, journalist Bridget M. Kuehn is ranked #17026, much ahead of 2021 Nobel Prize winner in medicine Ardem Patapoutian (ranked #28519), and 2022 Nobel Prize winner in physics Alain Aspect (ranked #20486).

Given that Goodhart's Law argues that "when a measure becomes a target, it ceases to be a good measure," this rank is susceptible to manipulation and abuse. Based on this standardized citation rank, John Ioannidis ranked himself as number 32.

In recent years, there has been a soar in the number of scientists who are listed as having published an excessive number of scientific articles, often at a rate much higher than their peers and many would find it implausible. While productivity is generally considered a positive outcome, researchers who are extremely productive run the risk of raising questions about the quantity, quality, and significance of their output. In some cases, cash incentives can be the driving force for hyper prolific authors. Ethically, it is questionable if these hyper-prolific authors even read their own papers, as they are often "honorary" authors who do not

deserve credit. It shows the danger of an over-reliance on publication data, some authors are mainly added to papers to complement the head of the lab with funding.

Hyper-prolific authors have a higher degree of citations because they generate more works that are encouraged for citation. Some use collaborative research strategy to boost papers, exposure, and editorial task. The conditioning of the data can maximize publications as journals are more likely to publish studies with conclusive, encouraging findings that were written by well-known, hyper-prolific authors. In short, being labeled as a hyper-prolific scientist should be cause for concern rather than pride. The "publish-or-perish" ethos that has dominated science for decades is the main reason behind their existence.

Table 5. Top most prolific authors in terms of number of papers and number of papers per year

| authfull | inst_name | np6021 | firstyr | lastyr | rank (ns) | h21 (ns) | np/year |
|---|---|---|---|---|---|---|---|
| Smith, George Davey | University of Bristol | 3791 | 1983 | 2021 | 98 | 198 | 99.8 |
| Lip, Gregory Y.H. | Aalborg University | 3484 | 1992 | 2022 | 208 | 146 | 116.1 |
| Wiwanitkit, Viroj | Dr. D. Y. Patil Vidyapeeth, Pune | 3366 | 1999 | 2022 | 30078 | 25 | 146.3 |
| Raoult, Didier | Aix Marseille Université | 3365 | 1979 | 2022 | 355 | 140 | 78.3 |
| Mahase, Elisabeth | BMJ | 661 | 2019 | 2022 | 33363 | 24 | 220.3 |
| Wiwanitkit, Viroj | Dr. D. Y. Patil Vidyapeeth, Pune | 3366 | 1999 | 2022 | 30078 | 25 | 146.3 |
| Iacobucci, Gareth | BMJ | 1364 | 2012 | 2022 | 95467 | 15 | 136.4 |
| McCarthy, Michael | Department of Seattle | 790 | 2012 | 2018 | 142232 | 13 | 131.7 |
| Rimmer, Abi | BMJ | 864 | 2014 | 2021 | 236367 | 11 | 123.4 |
| Ioannidis, John P.A. | Stanford University School of Medicine | 1264 | 1994 | 2022 | 32 | 157 | 45.1 |

Table 6. Top most prolific single authors. nps: number of single authored papers; rank (ns): nps/year: # single authored papers per year.

| authfull | inst_name | np6021 | firstyr | lastyr | rank (ns) | h21 (ns) | nps (ns) | nps/ |
|---|---|---|---|---|---|---|---|---|
| Dyer, Clare | BMJ | 2246 | 1984 | 2022 | 98067 | 13 | 2234 | 58.8 |
| Wiwanitkit, Viroj | Dr. D. Y. Patil Vidyapeeth, Pune | 3366 | 1999 | 2022 | 30078 | 25 | 1714 | 74.5 |
| Iacobucci, Gareth | BMJ | 1364 | 2012 | 2022 | 95467 | 15 | 1344 | 134.4 |
| Ernst, Edzard | University of Exeter | 2363 | 1980 | 2022 | 103 | 105 | 1312 | 31.2 |
| Kmietowicz, Zosia | BMJ | 1283 | 1996 | 2022 | 102462 | 15 | 1269 | 48.8 |
| Wise, Jacqui | Kent | 1231 | 2007 | 2022 | 80007 | 16 | 1228 | 81.9 |
| Kuehn, Bridget M. | | 989 | 2002 | 2022 | 17026 | 28 | 985 | 49.2 |

As an opposite, Table 7 presents the top most unprolific authors, who have just published 2 papers (from 1960) and yet in the top 2% list. A curious case is Irving Langmuir who first published in 1906 and last in 1997 (91 years) with only 2 papers but ranked #612. Irving Langmuir was an American chemist and physicist, who was awarded the Nobel Prize in Chemistry in 1932 for his work in surface chemistry. Another great example is G.H. Hardy who published from 1901 to 2003, yet only has 2 papers. Several bizarre entries again show some dubious and defective entries in the database.

Table 7. Top most unprolific authors

| authfull | inst_name | np6021 | firstyr | lastyr | rank (ns) | h21 (ns) |
|---|---|---|---|---|---|---|
| Langmuir, Irving | GE Global Research | 2 | 1906 | 1997 | 612 | 41 |
| Hardy, G. H. | Trinity College Cambridge | 2 | 1901 | 2003 | 21090 | 31 |
| Folch, Jordi | Harvard Medical School | 2 | 1939 | 1964 | 59924 | 9 |
| Grahame, David C. | Amherst College | 2 | 1937 | 1961 | 81014 | 16 |

| authfull | inst_name | np6021 | firstyr | lastyr | rank (ns) | h21 (ns) |
|---|---|---|---|---|---|---|
| Oppenheimer, J. Robert | Institute for Advanced Studies | 2 | 1926 | 1966 | 81517 | 17 |

## Most and Least cited authors

Citation is commonly used as the gold standard in citation metrics. Table 8 presents authors with the highest amount of citations, which are over 300,000 with an h index from 125 to 284. Author with the highest h index is Walter Willett who has 2251 papers since 1970 with an h index of 284.

Table 8. Top most cited authors, nc9621 (ns)*:* total cites 1996-2021 *(self citations excluded)*

| authfull | inst_name | np6021 | firstyr | lastyr | rank (ns) | nc9621 (ns) | h21 |
|---|---|---|---|---|---|---|---|
| Altman, Douglas | University of Oxford | 1046 | 1972 | 2020 | 11 | 428620 | 227 |
| Lander, Eric S. | Broad Institute | 643 | 1981 | 2022 | 38 | 372860 | 249 |
| Jemal, Ahmedin | American Cancer Society | 482 | 1993 | 2022 | 312 | 341530 | 125 |
| Grätzel, Michael | Ecole Polytechnique Fédérale de Lausanne | 1645 | 1971 | 2022 | 1 | 318947 | 252 |
| Willett, Walter C. | Harvard T.H. Chan School of Public Health | 2251 | 1970 | 2022 | 2 | 302008 | 284 |

As opposed to most cited authors, Table 9 lists top authors with the least amount of citations, with authors receiving less than 100 citations. This result is complemented in Table 10 with authors with the lowest h index of 3. Another interesting case is Joe H. Ward from Joint Base San Antonio a statistician, who has an h index of 3, but is ranked #86046 amongst researchers with an h index of over 30.

Table 9. Top least cited authors

| authfull | inst_name | np60 |
|---|---|---|
| Kund, Nirmal Kumar | VSS University of Technology | 44 |
| Nair, G. P. | Technocrat Society | 260 |
| Khan, Sameen Ahmed | Dhofar University | 96 |
| Torchigin, V. P. | Federal Research Center Informatics and Management of the Russian Academy of Sciences | 114 |

Table 10. Top authors with the lowest h index

| authfull | inst_name | np6 |
|---|---|---|
| Brown, Lisa Gottesfeld | Columbia University | 5 |
| Godard, Dominique N. | Compagnie IBM France | 5 |
| WARD, JOE H. | Joint Base San Antonio | 7 |
| Molodtsov, D. A. | Dorodnicyn Computing Centre of Federal Researcher Center Computer Science and Control | 7 |
| Frost, Otis Lamont | ARGOSystems, Inc. | 7 |

Self citation is authors who cite his/her own previously published work. Table 11 list authors with the most fantastic self citation of all time. G.P. Nair, Nirmal Kumar Kund and Oleg Mikahilov have over 92% self citation, while other have more than 90% self citation. It also displays that self citation can boost h index 3 to 5 times.

Table 11. Top authors with highest percentage of self citation. nc9621: total cites 1996-2021, Self%: Percent

self citation, h21: h-index as of end-2021; h21(ns): h21 self citation excluded.

| authfull | inst_name | np602 |
|---|---|---|
| Nair, G. P. | Technocrat Society | 260 |
| Kund, Nirmal Kumar | VSS University of Technology | 44 |
| Mikhailov, Oleg V. | Kazan National Research Technological University | 301 |
| Torchigin, V. P. | Federal Research Center Informatics and Management of the Russian Academy of Sciences | 114 |
| Piancastelli, Luca | Alma Mater Studiorum Università di Bologna | 107 |

## Authors who are an institute

Critical analysis of this database reveals that some authors in the list are an institution, such as Centers for Disease Control (CDC) and World Health Organization. CDC is ranked highly at #8520. Clearly, mixing a researcher who works hard with a centre of hundreds and thousands of researchers cannot be done in this so called standardized citation.

Table 12. Some of authors who are an institute

| authfull | inst_name |
|---|---|
| Centers for Disease Control | Centers for Disease Control |
| American College of Obstetricians | American College of Obstetricians |
| World Health Organization | Organisation Mondiale de la Santé |
| IARC Working Group on the Evaluation of Carcinogenic Risks to Humans | IARC Working Group on the Evaluation of Ca |
| Centers for Medicare | Centers for Medicare |

## Authors who are Journalists and Editors

Following on, many authors listed in the database are journalists and journal editors. And the papers published by these authors are not scientific papers but opinion pieces, editorials and news articles. They are not peer reviewed. Table 13 lists some of the authors. For example Richard Horton, the editor in chief of the Lancet has published 1388 papers and ranked #3167. Bridget Kuehn, a journalist, is ranked 17,026 higher than many highly regarded medical scientists and Nobel prize winners. Table 13 is just a list of some names, and it is unknown how many of these editors and journalists are in the database. Is it fair to compare a scientist who has to conduct many experiments to write a paper with an editor or journalist who can write an opinion daily? This standardized citation ranking is indeed questionable.

Table 13. Some of authors who are journalists and editors.

| authfull | inst_name | np6021 | firstyr | lastyr | rank (ns) | h21 (ns) | nps (ns) | nps/yr |
|---|---|---|---|---|---|---|---|---|
| Dyer, Clare | BMJ | 2246 | 1984 | 2022 | 98067 | 13 | 2234 | 58.8 |
| Iacobucci, Gareth | BMJ | 1364 | 2012 | 2022 | 95467 | 15 | 1344 | 134.4 |
| Kmietowicz, Zosia | BMJ | 1283 | 1996 | 2022 | 102462 | 15 | 1269 | 48.8 |
| Wise, Jacqui | Kent | 1231 | 2007 | 2022 | 80007 | 16 | 1228 | 81.9 |
| Kuehn, Bridget M. | Kuehn, Bridget M. | 989 | 2002 | 2022 | 17026 | 28 | 985 | 49.3 |
| Hawkes, Nigel | London | 865 | 2002 | 2019 | 184183 | 11 | 861 | 50.6 |
| Rimmer, Abi | BMJ | 864 | 2014 | 2021 | 236367 | 11 | 848 | 121.1 |
| Horton, Richard | The Lancet | 1388 | 1991 | 2022 | 3167 | 56 | 805 | 26.0 |
| Torjesen, Ingrid | BMJ | 783 | 2006 | 2022 | 169388 | 12 | 777 | 48.6 |
| O'Dowd, Adrian | BMJ | 762 | 2005 | 2022 | 309276 | 9 | 757 | 44.5 |

| authfull | inst_name | np6021 | firstyr | lastyr | rank (ns) | h21 (ns) | nps (ns) | nps/yr |
|---|---|---|---|---|---|---|---|---|
| Burki, Talha Khan | Burki, Talha Khan | 686 | 2008 | 2022 | 61453 | 19 | 685 | 48.9 |
| Mahase, Elisabeth | BMJ | 661 | 2019 | 2022 | 33363 | 24 | 653 | 217.7 |

# Conclusions

This study is the first of its kind to conduct a critical analysis of Stanford's list of the World's Top 2% Scientists. There are already several criticisms of such lists, which are frequently used as a metric for evaluating the impact and quality of a researcher's work. Among these criticisms are:

(1) The popularity of a particular field or research topic greatly influences the list. Regardless of the quality of their work, researchers working in highly cited fields are more likely to be included on these lists.

(2) The list does not consider the quality of citations received by a researcher. A researcher may receive a large number of citations, but if the majority of them are from low-impact or low-quality sources, they may not accurately reflect the researcher's work.

(3) The list does not include any early-career researchers. A young researcher may not have had enough time to accumulate a large number of citations, so the list may not give them the credit they deserve. Overall, the list of highly cited or top 2% researchers is a good indicator of impact, but it should not be used as the sole indicator of a researcher's quality and impact, as it has limitations and biases.

However, more importantly, this research reveals that the so-called standardized database of the world's top 2% of scientists is flawed. Among these flawed are:

(1) The database incorrectly listed researchers as first published in the nineteenth century and continued to publish until 2022.

(2) Many peculiar authors with low publication number and carrer lengths, for example, an author with only 2 papers but is ranked 612.

(3) Many authors with a large number of publications, and some of these are just news and editorial articles.

(4) Some of the authors listed in the database were journalists and editors, and their news articles were deemed "peer-reviewed" by this list.

(5) Some of the authors are an institute, not an individual

(6) Many authors with more than 50% self-citations in the list.

The study also discovered that there are deeply fundamental flaws in the so-called "databases of standardized citation indicators" which do not recognize if an author is a journalist and the articles are news articles. The use of such "standardized" ranking should not be encouraged.

# References

[1] Ioannidis, John PA, et al. "A standardized citation metrics author database annotated for scientific field." PLoS biology 17.8 (2019): e3000384.

[2] Ioannidis, J. P., Boyack, K. W., & Baas, J. (2020). Updated science-wide author databases of standardized citation indicators. PLoS biology, 18(10), e3000918.

[3] Ioannidis, John P.A. (2022), "September 2022 data-update for "Updated science-wide author databases of standardized citation indicators"", Mendeley Data, V5, doi: 10.17632/btchxktzyw.5

[4] Ioannidis, J., Klavans, R., & Boyack, K. W. (2018). Thousands of scientists publish a paper every five days. Nature 561, 167-169.

[5] Scanff, A., Naudet, F., Cristea, I.A., Moher, D., Bishop, D.V. and Locher, C., 2021. A survey of biomedical journals to detect editorial bias and nepotistic behavior. PLoS biology, 19(11), p.e3001133.