Scientometrics
https://doi.org/10.1007/s11192-023-04805-3

In search of a scientific elite: highly cited researchers (HCR) in France

Lauranne Chaignon[1] · **Domingo Docampo**[2] · **Daniel Egret**[1]

Received: 3 May 2023 / Accepted: 24 July 2023

Abstract
A high level of citations is generally associated with the exceptional influence of an article. Authors of several highly cited articles are thus considered and scrutinized as influential members within their scientific discipline and beyond when their activity spans several disciplines. Identifying individuals who have made outstanding contributions to science is the motivation behind the Highly Cited Researchers (HCR) list published annually by a commercial company. This article is devoted to the case of highly cited researchers affiliated with French institutions. We study the characteristics of the HCR population: productivity, gender bias, career and persistence on the list, collaboration network, and scientific integrity. Then, from the annual lists of HCRs from 2014 to 2022, we examine what the number of HCRs tells us about France's place in the world scientific arena, its evolution over a decade, and its geographical and institutional distribution. We discuss whether the population of French HCRs constitutes a scientific elite by examining the case of Mathematics. Finally, we discuss several indicators which could be used to verify that a researcher is close to the threshold to enter the list.

Keywords Highly cited researcher · Citation analysis · Bibliometric indicator · Research evaluation

Introduction

The search for excellence has received increasing attention in the scientific literature since the beginning of the 2000s (see, for example, van Leeuwen et al., 2003; Waltman, 2016). Most authors agree that defining excellence in a standardized and consistent way presents serious difficulties, especially since the definition can vary according to disciplines or research policies. In fact, they use a variety of terms to refer to excellence, such as "impact", "quality", "importance", "recognition", "influence", "value", "performance" and so on. It is sometimes difficult to say whether these terms are used in the literature

✉ Lauranne Chaignon
lauranne.chaignon@psl.eu

[1] Université PSL, 75006 Paris, France

[2] AtlanTTic Research Center for Telecommunication Technologies, Universidade de Vigo, Campus Universitario, 36310 Vigo, Spain

Published online: 14 August 2023

synonymously or whether they really reflect different concepts, especially since they may themselves have several meanings. Aksnes and Aagaard (2021) recall that the quality of a scientific article is a concept that covers several dimensions: “solidity, plausibility, originality, and scientific value”. In what follows, we have chosen, for the works we mention, to respect the terms used by their authors, thus avoiding distorting their intentions.

Citation indicators try and quantify research performance and are thus widely used to account for a publication’s impact and assess a researcher’s reputation (Martinez & Sá, 2020). On the other hand, although citations are widely recognized as an indicator of the importance of a published work (Szomszor et al., 2020), the link between citations and excellence or quality is worth discussing since many factors can influence the number of citations to an article (Aksnes & Aagaard, 2021; Bornmann & Daniel, 2008). There is also debate about how citation counts correlate with other forms of recognition (Basu, 2006; Warner, 2000). Articles distinguished by citations are not necessarily those that peers, or even the authors, judge to be the best (Aksnes & Aagaard, 2021; Borchardt & Hartings, 2018). Aksnes and Aagaard recall that Eugene Garfield himself, the founder of the citation indexes and one of the first to use citation analysis as a means "to identify and track creative people and their networks” (Docampo & Cram, 2019), considered that the frequency of citations was not in itself a sufficient indication to identify an exceptional, and influential publication.

However, despite all the fragilities mentioned above, citation indicators remain the most widely accepted quantitative tool to measure scientific impact. As summarized by Borchardt and Hartings (2018), citations, journal impact factor, and h-index constitute the triumvirate of impact evaluation. For nearly 20 years, we have witnessed the remarkable (and nevertheless controversial) success of the h-index, an indicator proposed by physicist Jorge Hirsch back in 2005; it has become the primary metric for quantifying the impact of an individual’s research (Koltun & Hafner, 2021). It is available from major bibliographic databases and has supported recruitment, promotion, or funding decisions. Despite its popularity, a relatively large number of shortcomings of the h-index have been identified (see, for instance, Teixeira da Silva & Dobránszki, 2018), prompting bibliometric practitioners to make use of normalized counts (Leydesdorff et al., 2016), as suggested in the Leiden Manifesto (Hicks et al., 2015). Bornmann and Leydesdorff (2018) indicate that once the number of citations to a paper is standardized according to the expected citation rate of the corresponding field of publication, a “reasonable alternative to the h-index is to count the papers which belong to the top-cited papers”.

Several studies have analyzed both the h-index and the authoring of highly cited papers (hereafter HCP) at the level of authors, institutions, or countries. Zhang et al., (2018), for Economics and Business, find a significant positive correlation and conclude that the HCP approach is useful for evaluating the impact of scholars, universities, or countries. Michael Schreiber (2013) highlights the arbitrariness of the thresholds used in both approaches and warns that the high threshold (Top1%) used for the HCP leads to an indicator that might not be robust enough because of the too limited number of papers.

Identifying highly cited researchers

There are different methods to identify highly cited researchers. For instance, Ioannidis et al. (2019) produced a database of the 100,000 most-cited authors across all scientific fields, using data from Scopus. However, the list by Clarivate™, based on the Web of Science and Essential Science Indicators™ (ESI), unarguably gets the most attraction, especially since the Shanghai ranking (ARWU) began using it to compose one of its indicators. In ARWU, the highly cited researchers indicator alone represents 20% of the total score, even though many universities ranked in ARWU present no HCR (Docampo et al., 2022). In this context, a single HCR can significantly change the positioning of a university in the ranking, a fact that has contributed to making universities attentive to this issue.

The number of HCPs constitutes the raw material for the list of highly cited researchers (hereafter HCR) released yearly by Clarivate to identify, based on their publications and the citations they have generated, the most influential researchers worldwide (Clarivate, 2022; Docampo & Cram, 2019). Thus, the list published in 2022 identifies approximately 4000 highly cited researchers in 21 fields of science and social sciences, and some 3200 additional highly cited researchers identified as being particularly influential in several of these fields (the so-called CrossField category). To build the list, Clarivate checks publications indexed in the Web of Science Core Collection™ over the previous decade (for the 2022 list: 11-year period 2011–2021). Highly cited papers (HCPs) are articles that rank in the first percentile (top 1%) in terms of citations for the field and year of publication. The document type is Article or Review (other types of documents are not considered). Highly cited researchers are those who satisfy two conditions: (1) whose total number of HCP lies above a certain threshold, adjusted in each category according to the size of the research field, as well as (2) a total number of citations to those HCPs enough to rank among all authors in the top 1% by total citations in the ESI field in which that researcher is considered.

To select the HCRs, Clarivate distributes the articles and reviews among 21 disciplines according to the journal in which they appear. In the case of multidisciplinary journals (for instance, Science, Nature, PNAS, PLOS), Clarivate attributes the documents to a single discipline after analyzing their content.

To count citations, Clarivate uses a whole count (i.e., each co-author of a publication receives the total citation count), not a fraction based on the number of co-publishers. Note that this approach is discussed by authors who advocate the use of fractional counting, especially in the case of the evaluation of scientific entities (see Waltman, 2016 for a summary of this discussion).

Within each ESI category, Clarivate estimates the number of HCRs retained from the square root of the number of authors appearing in the discipline. The researchers, ranked in descending order of HCP, are included in the list until the number of HCP reaches a threshold to accommodate the precomputed quota of eligible candidates. The list also includes authors who present only one less HCP publication than the last HCR selected, provided that the total number of citations of their HCPs lies in the upper half of those obtained by the HCP articles of the HCRs already present on the list. A researcher may appear several times in the list when the publications fall within several disciplinary fields or otherwise appear in the "cross-field" category, which considers researchers whose number of HCPs lies below the threshold in some disciplines but whose cumulative number of citations is equivalent to or greater than that of the researchers selected in the ESI categories.

The methodology first employed by Thomson and then by Clarivate to develop this list of HCRs has evolved (Docampo & Cram, 2019). In the first release of 2001, the list only accounted for the number of citations obtained by a researcher over a given period and disciplinary field. In 2012, Clarivate modified the methodology by introducing the number of HCPs in the calculation. A second change took place in 2018, with the appearance of the "cross-field" category. Finally, Clarivate has introduced restrictions to including publications with many co-authors (i.e., articles with more than 30 authors) or the withdrawal of potential HCRs with too high a level of self-citations.

Looking for pioneering researchers

The stated objective of the promoters of this list is to identify the individuals (and, beyond these individuals, the institutions, and laboratories) who are "true pioneers in their fields"; i.e., who have made outstanding contributions to the development of science and technology over the past decade. The scientists identified in these lists are supposed to be “individual geniuses […] with extraordinary contributions” (Aksnes & Aagaard, 2021) with research achievements that will show scholar influence as well as a substantial impact on society and the economy (Nagane et al., 2018).

According to several studies, the number of highly cited papers produced during a given period is also a relevant indicator to characterize the production of a laboratory, an institution, or a country (Bornmann, 2014; Tijssen et al., 2002).

As a result, the analysis of the HCR list gives rise to two main approaches: the highly cited researchers as a population and as an indicator.

- As a population, the subject of analysis is the characteristics of the highly cited researchers, their academic age, and their research practices.
- As an indicator, the number of highly cited researchers attracts interest to measure the impact, even the excellence of research, and to make comparisons between research fields, institutions, or countries.

This paper is devoted to the specific case of highly cited researchers affiliated with France. First, we study the characteristics of the HCR population: productivity, gender bias, career and persistence on the list, collaboration network, and scientific integrity (see Sect. "Highly cited researchers as a population: portrait of the French HCRs"); then we examine what the annual lists of HCRs from 2014 to 2022 provide as information on the place of France in the world scientific arena (see Sect. "Using the number of highly cited researchers as an indicator"). Finally, we discuss in Sect. "Discussion" to what extent the French HCRs can be considered as a scientific elite, and we propose a new indicator for identifying potential HCRs.

Highly cited researchers as a population: portrait of the French HCRs

We describe in this section the composition of the population of French HCRs, paying particular attention to their scientific output, the breakdown by gender, their age and length of presence on the lists, the degree of collaboration between the HCRs, their other types of recognition, and the question of integrity.

The scientific output of HCR

General characteristics

First and foremost, HCRs are very prolific researchers (Aksnes & Aagaard, 2021; Must, 2020). Studying African HCRs, Confraria et al. (2018) found that, on average, they produce three times more publications per year than non-HCRs. They note that researchers who have published most articles usually have a higher reputation and swiftly get the resources to facilitate research. The correlation between productivity and citation leads to identifying productivity as a determining factor for being among the highly cited researchers. On the other hand, HCRs can be prolific to the point of raising questions; for instance, some HCRs affiliated with King Abdulaziz University publish an article every other day, hardly representative of the research activity of an individual (Alhuthali & Sayed, 2022).

However, a recent study on the ability of researchers to produce highly cited publications found that productivity alone is not the only mark of influential scientists, since it cannot be easily disembodied from luck when analyzing the scientific impact of a paper. Sinatra et al. (2016) proposed the Q-model, a bibliometric measure that would unlock the secret behind scientific success by untangling the role of productivity, luck, and individual ability in a scientific career. We will come back, in the Discussion section, to the Q factor and to other indicators that may be used for detecting and explaining extremely high productivity.

Regarding the sources of publications, it is worth noticing that many HCPs are published in journals with high-impact factors, such as *Nature*, *Science*, or *PNAS* (Aksnes & Aagaard, 2021; Sinay et al., 2020).

To complete this general picture of HCR publication practices, we can point out that a significant part (about 28%, for the current decade) of their highly cited publications with less than 30 authors are not articles, but reviews, which is not unexpected since reviews usually serve as references of the state of the art for a given research question, and thus generate many citations (Aksnes & Aagaard, 2021). The authors of this study found that 7% of all HCRs mainly or only published highly cited review articles.

At the international level, the share of reviews in the total of HCPs varies significantly depending on the country, reaching 41% for India and falling to 19% for China (source: InCites™).

HCPs are more frequently in Open Access (March 2023: 57%) than the world reference for articles and reviews of the same time span 2011–2021 (45%).

French HCRs

Let us now focus on the HCPs for the French HCRs from the latest annual list. For the present study, we analyzed a corpus of 1871 highly cited papers from the ESI list, published in the decade 2011–2021, for the 136 HCRs affiliated in France.

Regarding the journals where French HCRs have published the HCPs, we find major medical journals (e.g., oncology, cardiology) alongside notorious multidisciplinary journals such as Nature, Science, and PNAS (see Table 1). Over 95% of HCPs appeared in first-quartile journals. All HCPs, with one exception, are written in English.

As already observed, the reviews represent a large part of the highly cited papers: the share of reviews for France is 24%, whereas the overall share of reviews for all papers (articles and reviews) with an affiliation in France, regardless of their citation rate, is only 7%. In addition, this share varies greatly from one discipline to another: 56% of the HCP in Chemistry are *review papers*, but only 5% in the domain of Space Science, and 2% in Mathematics.

Regarding the ability to disseminate and promote the results of their research, we observe that 72% of the 1871 HCPs (France-HCP 2022 list, articles from 2011 to 2021) are in Open Access (March 2023), while the world reference for the HCPs in the same period is 57%. The baseline of share of Open Access for articles and reviews published with at least a French affiliation is 61%.

Table 1 Journals having published at least 20 HCPs of French HCRs (2011–2021)

Source	# HCP
NEW ENGLAND JOURNAL OF MEDICINE	87
LANCET ONCOLOGY	72
JOURNAL OF CLINICAL ONCOLOGY	65
NATURE	56
LANCET	47
NATURE COMMUNICATIONS	45
ANNALS OF ONCOLOGY	42
SCIENCE	30
PNAS	30
EUROPEAN HEART JOURNAL	24
GUT	24
NATURE MATERIALS	22
CLINICAL CANCER RESEARCH	21
GASTROENTEROLOGY	20
EUROPEAN JOURNAL OF CANCER	20

Breakdown of HCRs by gender

General characteristics

The gender of HCRs is a well-studied characteristic (Bornmann et al., 2015; Meho, 2022). In his comprehensive study, Meho considers all HCRs from 2014 to 2021 and finds that women are even more underrepresented among the elites than in the rest of the scientific community, representing only between 13 and 14% of HCRs in the period. Among the sources of this inequality, it is worth mentioning the lesser place given to women in scientific project management functions and the observation that collaborative networks are less international for female scientists than for their male colleagues Other researchers conduct gender-based analyses of HCR at a national or regional scale. For instance, in China and Germany the proportion of women among HCRs does not exceed 5% or 6%, respectively (Bornmann et al., 2017; Wei & He, 2021). Sinay et al. (2020) also reported the absence of female researchers from Latin America, Africa, Asia, and Oceania.

The presence of women in the HCR lists varies not only with the country of affiliation but also with the disciplines. It is in the social sciences that women are most numerous (31%), whereas their proportion is the lowest in engineering (4%) (Bornmann et al., 2015). Meho (2022) notes that, for STEM, the gender gap among HCRs peaks in chemistry, computer science, engineering, mathematics and physics, and astronomy. Things began to change after the introduction of the Cross-Field category in 2018 since it is in this category that the share of women has increased the most (Shamsi, 2021).

French HCRs

In France, the share of women among HCRs between 2014 and 2022 is 14.8%, slightly higher than the world average of around 13.4% (see Table 2 and Meho, 2022). For reference, the official national report on the state of Scientific Employment in France (MESRI,

Table 2 Number of HCRs and proportion of women, in France and worldwide, by year

	2014	2015	2016	2017	2018	2019	2020	2021	2022	Individuals
Total individuals All	3215	3126	3266	3538	6079	6216	6389	6602	7225	45,656
%Women (All)	12.8%	13.2%	13.4%	13.0%	13.7%	13.5%	13.5%	13.6%		13.4%
Total individuals France	83	72	97	89	157	156	161	147	136	1098
%Women (France)	9.6%	11.1%	13.4%	14.6%	15.9%	14.7%	16.8%	15.6%	16.2%	14.8%

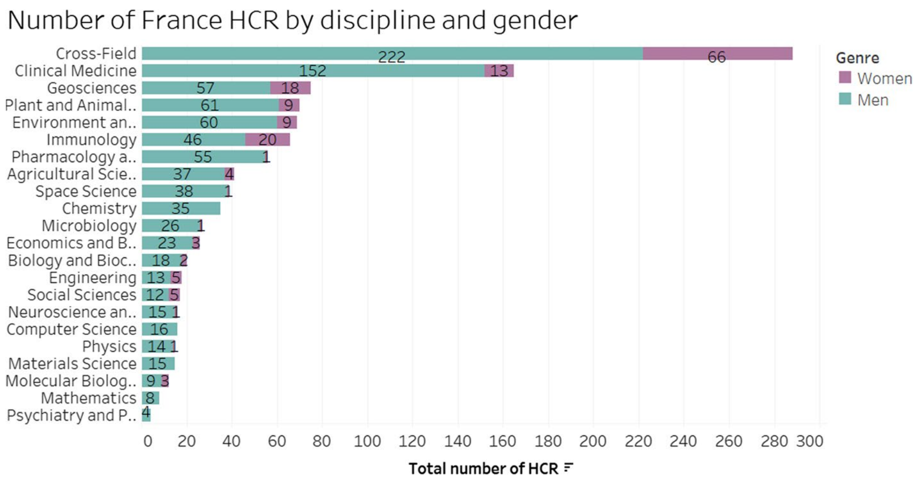


Fig. 1 Number of France HCR by discipline and gender. © Tableau software

2023) mentions a share of women of 40.6% in French public research institutions in 2020. However, this proportion is only 29% among research and academic staff of grade A (grade of Professor).

This gender bias is in line with the results of the study by Larivière et al. (2013), which show that the gender inequality observed in the citations is particularly notable for the most productive countries of scientific publications.

Figure 1 shows the breakdown of French HCR by gender and discipline. It is worth highlighting the contrast between Cross-Field and Clinical Medicine. In the case of chemistry, computer science, mathematics, psychiatry/psychology, and materials science there is no French female researcher listed as HCR between 2014 and 2022.

Career length

General characteristics

Another characteristic that deserves attention is the age of the HCRs and the length of their career. Academic seniority does not seem to affect the impact of an article: it is the quality, novelty, and interest of the scientific community in the subject that takes precedence (Must, 2020). Both young and senior researchers can be the authors of highly cited articles. However, the average age of German HCRs is 55, which corresponds, within 2 years, to the average age at which German professors are recruited in this country (Bornmann et al., 2017). 85% of German HCRs have this status. Most highly cited researchers have been in their field for 11 to 15 years (Must, 2020). Similarly, more than half of Chinese HCRs first appeared on a list during the so-called “development” period of their career, i.e., between 11 and 20 years after their debut (Wei & He, 2021). In complement, they mention that the productivity of Chinese HCRs increases over the course of their career, peaking between 30 and 36 years after the publication of their first paper.

Table 3 Average academic age (years after Ph.D.) of HCRs affiliated in France on the Clarivate lists from 2014 to 2022 by discipline

Discipline	Average academic age
Molecular Biology and Genetics	32.0
Pharmacology and Toxicology	31.4
Biology and Biochemistry	31.0
Physics	28.8
Mathematics	28.2
Engineering	28.2
Chemistry	26.3
Clinical Medicine	26.0
Psychiatry and Psychology	26.0
Materials Science	24.0
Immunology	23.9
Geosciences	23.6
Cross-Field	22.3
Neuroscience and Behavior	21.6
Economics and Business	21.4
Computer Science	21.0
Microbiology	20.5
Plant and Animal Science	19.8
Agricultural Sciences	18.9
Environment and Ecology	18.7
Social Sciences	16.7
Space Science	16.3
Total	23.2

By and large, HCRs stay on the list between 1 and 5 years, with an average of two and a median of three, a sign of the noticeable volatility of the list of highly cited researchers.[1]

French HCRs

In the following, we consider the academic age of French HCR, calculated from the year of their Ph.D. thesis. Of the 351 individuals who appear with a French affiliation in at least one of the Clarivate lists (2014 to 2022), we collected data about the Ph.D. year of 340. Of the 11 HCRs whose data could not be collected, seven do not seem to hold this degree (they are engineers).

The average (and median) academic age of the 340 HCRs is 23.2 years at their first appearance on the list (Table 3). The average age at which a Ph.D. is obtained is 30.7 years in France (MESRI, 2023). This implies that the average age of French HCR, at their first

[1] In the Medical disciplines, for French physicians, we did not use the year of their thesis defense but rather the year of obtaining a doctorate, which may push back the average academic age for this category by a few years. It should be noted that for studies calculating the academic age from the year of first publication, there could be a difference of one or two years with our calculations based on the year of Ph.D., the first publication most often occurring before Graduation.

nomination on the Clarivate list, is around 54 years, very similar to that observed by Bornmann et al. (2017) for German researchers and could also be compatible with the "development" period referred to by Must (2020) and (Wei & He, 2021).

We found that almost a quarter of the HCRs affiliated in France obtained their doctorate abroad, primarily in the United Kingdom (13 HCR), the USA (10 HCR), Germany, and Italy (9 HCR). It would be interesting to analyze the impact of this foreign experience on the careers of French HCRs, particularly on the networks of collaborations that these authors benefit from today.

Collaborative networks (national and international)

General characteristics

Although it runs against the perception of HCRs as exceptional individuals, HCPs are usually multi-coauthored. The article, not the individual researcher, is the target of citations; publications, particularly highly cited ones, are generally the fruit of collective work (Aksnes & Aagaard, 2021). This type of scientific collaboration can take place at different scales. Confraria et al. (2018) found that South African HCRs frequently collaborate with other researchers, whether within their institution, with other institutions in the country, with other African countries, or with institutions on different continents. North America and Western Europe are at the heart of global collaboration networks (Martinez & Sá, 2020), and in general, the HCRs remain loyal to their collaboration partners since these usually last between 6 and 10 years (Must, 2020). The interest of the authors of articles with many coauthors comes from the widely accepted positive effects of collaborations on the impact of research (Confraria et al., 2018). Martinez and Sá (2020) also found that publications from medium- and large-scale international teams receive many citations; it is thus no surprise that most Brazilian HCPs result from large-scale collaborations.

It is important to note, in this context, that Clarivate has chosen in recent years to exclude from its analysis publications with more than 30 authors, or signed by an explicit group of authors, arguing that it would be unreasonable to credit one author out of several dozens or even hundreds. This decision may favor theoretical papers and review papers potentially signed by a smaller number of coauthors than papers describing, e.g., complex instrumentation. This restriction has a significant impact on certain disciplines, such as Physics and Space science, which lose 29% and 55% of their HCPs affiliated with France between 2011 and 2021, respectively.

French HCRs

Highly cited papers by French HCRs, as illustrated in Fig. 2, present a distribution of the number of authors with a noticeable degree of concentration in the range 2–7, followed by a flat tail for large numbers of coauthors.

85% of the articles resulted from international collaboration. The reference for France (all articles and reviews, 2011–2021) is 61%. Table 4 shows that HCRs from France co-authored 52% of their 1871 HCPs with researchers affiliated with institutions in the United States. The international collaborations of French HCRs spread worldwide, although Europe and North America appear to be the largest sources of partnerships (Table 4). Note that the frequency distribution of other-country collaborations in French HCPs partly

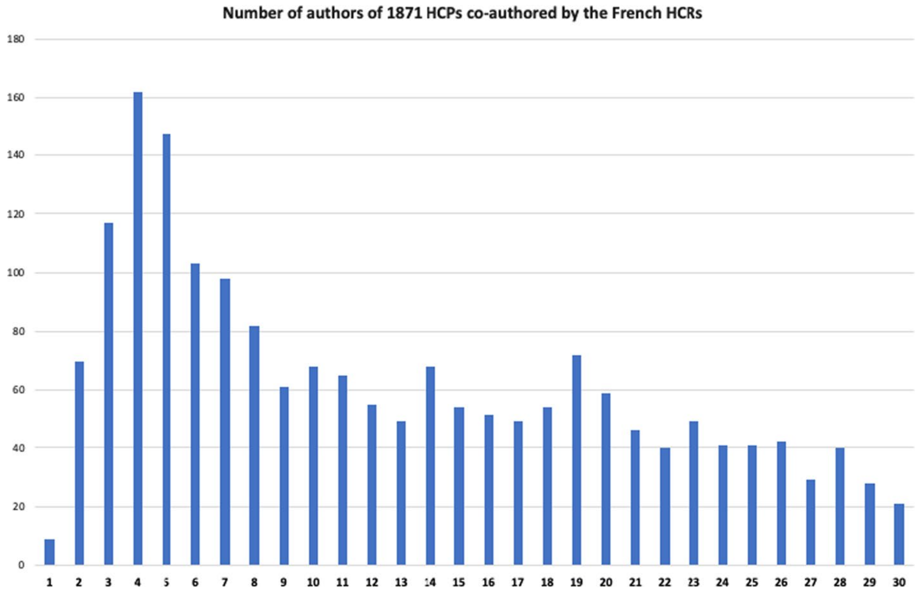


Fig. 2 Number of coauthors of highly cited papers by French HCRs. The figure only refers to articles with less than 30 authors

Table 4 Numbers and shares of the HCP articles in an international collaboration by country (only share values higher than 10% are listed)

Country	# HCP	%
France	1784	95
USA	976	52
United Kingdom	624	33
Germany	522	28
Italy	399	21
Spain	325	17
Netherlands	296	16
Australia	295	16
Canada	273	15
Switzerland	231	12
China	234	13
Belgium	215	11

Corpus: 1871 HCPs (2011–2022) co-authored by HCRs affiliated in France

reflects the availability of potential collaborators (e.g., notably larger with the USA than with the UK) as well as other underlying factors.

HCRs also collaborate among themselves. Figure 3 is a visualization of the collaborations (or lack of) between the HCRs affiliated with France on the 2022 Clarivate list. The darkest circles, connected by lines, indicate one or more co-publications between HCRs. Conversely, the lighter and isolated points show that these HCRs did not publish an article in collaboration with another HCR from the 2022 list from 2011 to 2021. It turns out

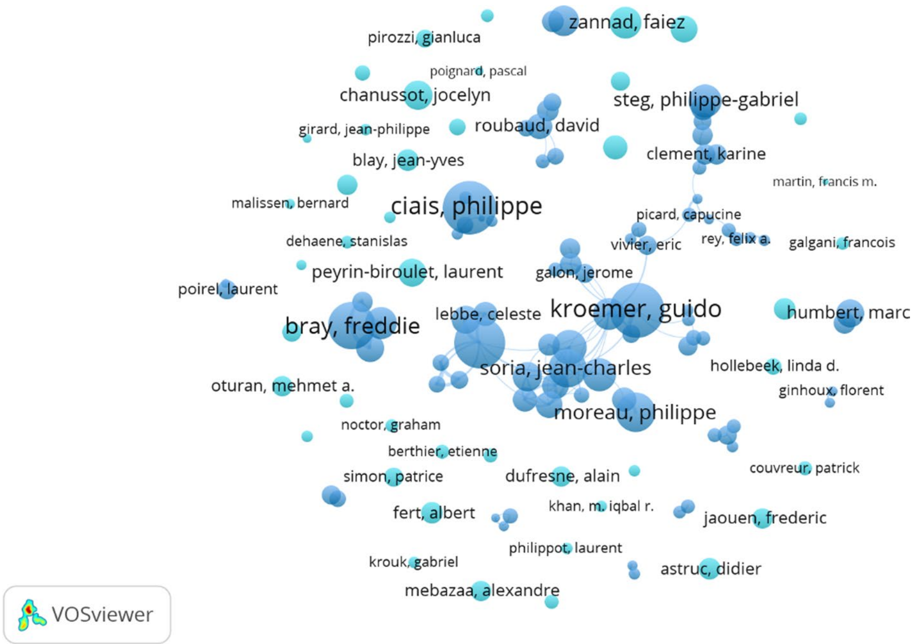


Fig. 3 Visualization of collaborations between French HCRs

that about 66% of the HCRs on this list have at least one collaboration with another HCR during this period. These HCRs co-authored HCPs with 2.8 other HCRs on the list on average, with a maximum of 11. The disciplines most concerned by these co-publications between HCRs lie in the health sciences (Clinical Medicine, Immunology, Pharmacology) and cross-field category. Figure 3 includes HCR interconnects that have arisen due to (1) the co-publication of the actual HCPs that have made them both HCRs, as well as (2) other factors (such as long-standing collaborations between their institutions).

HCR vs other forms of recognition

General characteristics

The work of American sociologists of science first sparked academic interest in scientific elites. Harriet Zuckerman (1977), in her study of Nobel laureates, defined a scientific elite as the group of individuals who have made a difference in advancing scientific knowledge. She called the Nobel laureates an ultra-elite and frequently used membership in the American Academy of Sciences (NAS) as an elite scientific benchmark.

French HCRs

It is thus worth checking the membership in the French Academy of Sciences of French HCRs. The Academy comprises 285 members, 119 foreign associates, and 62 correspondents elected among the most eminent French and foreign scientists. The overlapping of the two lists—Academicians and the 351 HCRs from Clarivate lists 2014 to 2022—is short since only 21 French HCRs (6%) are members of the Academy. This special recognition by peers does not seem to go along with the performance in the number of citations. However, this comparison has some limitations: on the one hand, the humanities, economy, social sciences, and medicine are not present in the French Academy of Sciences; on the other hand, the average age of academicians is significantly higher because they are elected for life and, therefore, only partly coincides with the age range of HCRs, a side-effect of the 11-year window for which influential publications count.

The outcome is similar when comparing the list of HCRs with the list of CNRS Silver and Gold medals recipients. These prizes, awarded yearly following recommendations from the French National Committee for Scientific Research, distinguish individuals working in French laboratories for "the originality, quality, and importance of their work". Of the 49 winners of the CNRS Gold Medal since 1979, only 7 (i.e., 14%) have been HCRs. Regarding the 414 Silver Medals awarded since 2001 (about twenty each year in different scientific fields), only 24 winners (i.e., 6%) are also present in the lists of the French HCRs.

We can also mention the case of the 10 French Nobel Prizes winners in Physics, Chemistry, or Medicine in the last two decades, who were not highly cited authors except for Albert Fert (Physics, Nobel 2007, who entered the HCR list in 2018).

The noted misadjustment is hardly a French phenomenon. At the end of the 1960s, E. Garfield tried to identify researchers being “of the Nobel class” thanks to citation analyses (Aksnes & Aagaard, 2021; Docampo & Cram, 2014). He considered the Nobels of the 1980s to be notable for their extraordinary abilities to produce highly cited papers, although those scholars were otherwise not particularly productive. However, this indicator as a prediction tool no longer seems valid since 82% of the Nobel laureates of the last decade were not HCR, and 37% had not published any HCP (Kosmulski, 2020). The effectiveness of scientometric measures appears to be declining due to the rise of hyper authorship (Koltun & Hafner, 2021). Moreover, less than 25% of highly cited researchers got any scientific prize (Koltun & Hafner, 2021).

As of the 136 French HCRs, 16 have received one or more research prizes included in the list of prestigious international research prizes established by Meho (2020). Conversely, among the 93 French prize winners in the last two decades (2001–2020), only 16 (i.e., 17%) are currently or were previously highly cited researchers, the other 77 were not part of Clarivate’s HCR lists.

In another register, ERC, i.e., grants from the European Research Council, confer a label of excellence to the selected researchers. 60 of the 351 researchers identified in the Clarivate lists (17%) obtained at least one ERC grant. Of these 60 researchers:

- 47 HCRs got their first ERC before appearing for the first time in these lists (on average, 5 years before).
- 7 HCRs appeared for the first time on a list before having their first ERC (on average, 4 years before making it to the list).
- 6 HCRs had their first ERC and made the HCR list for the first time the same year.

It seems clear that a convergence of causes (multi-authorship, increasing pressure on the citation process) contributes to blurring the natural correlation between the repeated production of highly cited articles (which can legitimately be considered particularly influential) and the recognition by peers of singular merits through medals and prizes. Although this should not question the academic prestige of these influential researchers, it nevertheless leads to caution in interpreting the HCR indicator as an index of excellence for an institution or a discipline.

Scientific integrity: self-citations

General characteristics

HCR self-citation practices vary across disciplines, countries, and the stage of a researcher's career (Szomszor et al., 2020; Van Noorden & Chawla, 2019). For Szomszor et al. (2020), the function of a self-citation is the same as a classic citation: to indicate the publications on which a new work depends, the links with it, and the positioning of the author. They are part of the fabric of science and should not be excluded from citation indicators and the resulting analyses. However, this practice can also be a tactical tool to gain visibility, thus distorting the assessment of an individual's search performance. Although defining what constitutes an excess of citations is tricky, Szomszor et al. (2020) estimate that the standard self-citation rate was around 8% in the 1960s and has changed little in 50 years. With regard to this rate, these authors have analyzed the practices of the HCRs and found them within the normal range of average self-citations, regardless of their scientific field, with just one exception. Indeed, self-citations in mathematics may exceed 30%, but this seems to be associated with the existence of small, isolated subfields.

All the same, there are authors among the HCRs whose practices raise questions. Particularly when self-citations, whether from the author himself or the co-authors, represent more than 70% of the citations obtained by an article (Van Noorden & Chawla, 2019). Indeed, an American biophysicist was removed from the lists in 2019, at the initiative of Clarivate, due to an unusual number of self-citations. Moreover, other accusations of citation manipulation weigh against this researcher, practices that seem to extend over several decades. (Van Noorden, 2020).

French HCRs

For French authors (not only HCRs but all authors) of Highly Cited Papers, the self-citing average, according to InCites™, is 4%. We find the same average of 4% for the 136 French HCRs nominated in 2022. The maximum value found is, for this sample, 20%.

Clarivate mentions the desire to exclude authors with a very high rate of self-citation, the threshold being established on the basis of deviations from the median within each ESI discipline, but one can assume that this 'very high rate' concerns exceptional cases where rates are significantly higher than those observed here (in any case, well above 30%). We can safely conclude that self-citation practices are not an issue in the case of the French HCR population.

Scientific integrity: the issue of retracted publications

General characteristics

Self-citations are not the only issue of interest concerning scientific integrity. The participation of HCRs in retracted publications is another matter of concern. For instance, in the context of increasing numbers of publications in Iran, Kamali et al. (2022) found that 10% of Iranian HCRs between 2006 and 2019 had at least one of their articles retracted. In two-thirds of the cases, articles were retracted for duplication and false peer reviews. The authors of this study consider that these misconducts should preclude an author from appearing in the Clarivate lists.

From 2022, selected HCR authors must go through a publication ethics filter based on the *Retraction Watch* database (http://retractiondatabase.org/), which Clarivate uses to identify all authors who have had one of their articles retracted or contested. More precisely, in its methodology, Clarivate states that "Beyond this, researchers found to have committed scientific misconduct in formal proceedings conducted by a researcher's institution, a government agency, a funder or a publisher are excluded from our list of Highly Cited Researchers."

Many of these integrity issues occur in life and health sciences. The announced rate of authors rejected by Clarivate, 7% or 550 out of 6938, is very noticeable.

French HCRs

Given the strong representation of French HCRs in the life and health sciences, we believe the rate for France is likely somewhat higher, close to 10%. Based upon computations following Clarivate's guidelines for the nomination of HCRs in 2022, we estimate that about 15 French researchers, some of whom were HCRs in previous years, were excluded from the 2022 list (i.e., 10% of some 150 potential HCRs), some of them in the aftermath of the controversies related to the COVID-19 pandemic.

Using the number of highly cited researchers as an indicator

Beyond individuals, HCR counts can be used to gain insight into the research performance of entities and countries, as well as the dynamics at a national or international level. For example, they can help identify trends in global leadership in science and technology (Basu et al., 2018). In 15 years, China has thus become one of the largest providers of HCR in the world. The remarkable increase in high-impact researchers undoubtedly owes to the country's advance in science and technology (Basu et al., 2018; Li, 2018; Wei & He, 2021). Despite the apparent progress of China, the world landscape of HCRs remains dominated by the USA and, in general, by Western countries. However, a slight decrease is recorded in the USA, China, and Japan (Nagane et al., 2018; Nielsen & Andersen, 2021). The number of HCRs can be related to the scientific production of a country, as in the citation excellence indicator developed by Basu in 2006.

Position of France in the ranking and evolution

In the annual list published in November 2022, France ranks eighth in the world for the number of highly cited researchers (Table 5). We can put this rank in perspective with the seventh world rank of France for the count of the total number of publications (articles and reviews) in the period 2011–2021. As Table 5 shows, Australia or the Netherlands home more HCRs than France, while their production of articles is lower (significantly lower in the case of the Netherlands).

The eighth place worldwide in the number of HCRs has remained stable since 2016 (see Table 6). On the other hand, the world's share of French HCRs is down, consistent with the simultaneous decline in the proportion of French scientific contributions to the global production of articles and reviews.

Table 5 Number of HCRs and articles referenced in the Web of Science (2011–2021), ranked by country. Data for France are in bold

Country	Nb HCR	World Rank HCR	Nb articles & reviews (11 years)	World Rank publications
United States	2764	1	5,058,887	1
China Mainland	1169	2	4,022,460	2
United Kingdom	579	3	1,342,310	4
Germany	369	4	1,347,655	3
Australia	337	5	805,079	10
Canada	226	6	873,000	9
Netherlands	210	7	500,872	15
France	**136**	**8**	**910,754**	**7**
Saudi Arabia	114	9	210,999	24
Switzerland	112	10	379,413	18
Singapore	106	11	166,176	34
Italy	104	12	889,946	8

Table 6 Number of French HCRs and articles referenced in the Web of Science (2011–2021), world rank, and global share by year

Year of the list	Nb HCR France	World Rank	% France/ World	Nb articles & reviews (11 years)	World Rank	% France/ World
2014	83	7	2.6	679,245	6	5.3
2015	72	9	2.3	701,161	6	5.3
2016	97	8	0.03	724,425	6	5.2
2017	89	8	2.5	754,427	6	5.1
2018	157	8	2.6	783,460	6	4.9
2019	156		2.5	811,894	6	4.8
2020	161	8	2.5	840,955	6	4.7
2021	147	8	2.2	874,561	6	4.5
2022	136	8	1.9	910,754	7	4.4

Rankings of countries based on the number of HCRs have the disadvantage of being size-dependent. To overcome this limitation and to highlight the specific characteristics of research quality, Alonso Rodríguez-Navarro and Brito (2022) suggest normalizing the number of HCRs by the number of inhabitants, as well as by the gross domestic product (GDP). The two approaches result in different rankings but with a common logic: the largest countries that are at the top of the rankings in terms of the number of HCRs see their position drop when this number is normalized, in particular China.

The calculations of Rodríguez-Navarro and Brito (2022) were based on the highly cited researchers from the work of Ioannidis et al (2019). Carrying out this normalization by GDP on the 2022 list of HCRs from Clarivate, we can also observe the resulting changes in the rankings: USA and China fall back, as does France. In Table 7, we compare France's ratio HCR/GDP with other countries sharing similar sizes and values

Table 7 World rank by number of HCR per billion USD of GDP

Country	Nb HCR 2022	Average GDP 2011–2021	Ratio HCR/GDP (%)	Ranking
United Kingdom	579	2855.65	200	1
Saudi Arabia	114	732.28	160	2
Canada	226	1737.17	130	3
Germany	369	3766.60	100	4
Spain	97	1344.97	70	5
France	136	2712.68	50	=6
Korea	70	1527.95	50	=6
Italy	104	2041.69	50	=6
Japan	90	5192.70	20	9

In this table, we consider only countries with a population between one-half and two times the one of France, and with a GDP per capita between one-half and two times France. A ratio of 100 means that the number of HCRs is commensurate with GDP. Departures from 100% indicate a surplus or a shortage of HCRs. *Source of the data:* World Economic Outlook Database, October 2022

of GDP per capita (which is known to better represent the socioeconomic development stage of a country).

We must highlight that if we perform the same exercise of normalization by GDP based on the number of Nobel Prizes awarded between 2010 and 2022, France occupies the 5th place worldwide (fourth in terms of the number of prizes).

Dynamics of research fields through HCR counts

The analysis of the distribution of HCRs by discipline sheds light on the fields of research specialization in the country.

Table 8 below shows the percentage of French HCRs, by ESI category and year. We note that the most represented categories, for France, compared to the world, are Clinical Medicine and the other disciplines associated with the Health Sciences (Immunology, Microbiology, Pharmacology & Toxicology), as well as Economics and Geosciences. However, the latter appears to be in sharp decline compared to a peak reached in 2015–2017, as also happens with Space Science.

Table 9 highlights the increase in the number of French HCRs in 2018, with the appearance of the Cross-Field category (an overall increase of around 70% in the number of HCRs), with no changes nonetheless in the global share: 4.9% in the 2017 list for 4.8% in the 2018 list. The table also shows a decline in the total number of HCRs in 2021 and 2022. It is worth reminding that some researchers are identified and thus counted as highly cited in several scientific categories. Hence, the total number of HCRs exceeds the total number of individuals included in the list: for instance, the 136 HCRs affiliated in France in 2022 correspond to 132 individuals.

For France, the three disciplines with the most HCRs in cumulative number over the 9 years shown in Table 9 (apart from the Cross-Field category) are clinical medicine (165, showing consistent growth over the 9 years), geosciences (75), and plant and animal science (70).

The three disciplines with the fewest HCRs are psychiatry & psychology (4), mathematics (8), molecular biology, and genetics (12). The absence of French HCRs in mathematics

Table 8 Distribution by ESI fields of research and year of the world's share of HCRs affiliated with French institutions. The darkest values in the table are the highest. (Color figure online)

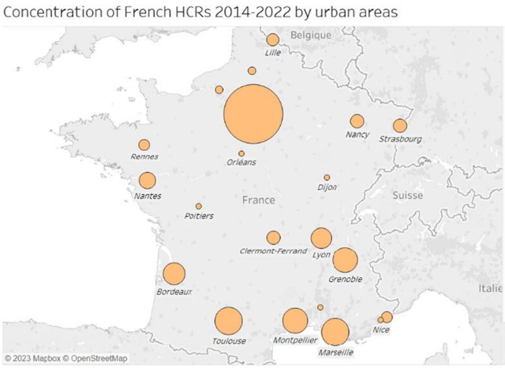
Areas of specialisation of French HCR

	2014	2015	2016	2017	2018	2019	2020	2021	2022	Total
Cross-Field					2,7%	2,4%	2,4%	2,1%	1.7%	2.2%
Agricultural Sciences	4,5%	3,9%	5,3%	5,8%	2,5%	2,7%	3,6%	2,4%	0.0%	3.5%
Biology & Biochemistry	1,0%	1,3%	1,5%	0,9%	1,6%	1,9%	0,8%	0,0%	0.0%	1.0%
Chemistry	1,5%	2,9%	3,2%	2,8%	1,9%	1,7%	0,8%	0,4%	0.4%	1.7%
Clinical Medicine	3,2%	2,9%	3,4%	3,1%	4,2%	4,6%	5,0%	5,7%	5.4%	4.3%
Computer Science	1,7%	0,9%	1,6%	1,3%	2,1%	1,9%	1,6%	0,9%	1.7%	1.5%
Economics & Business	2,1%	2,8%	2,9%	3,2%	4,2%	2,7%	3,0%	3,7%	4.3%	3.2%
Engineering	4,8%	1,3%	2,1%	0,6%	0,0%	0,0%	0,0%	0,6%	1.3%	1.2%
Environment & Ecology	6,6%	6,1%	6,8%	3,8%	4,3%	3,6%	4,0%	4,0%	3.0%	4.5%
Geosciences	5,7%	6,8%	7,4%	6,4%	4,9%	5,2%	5,3%	4,2%	3.4%	5.5%
Immunology	0,0%	2,7%	7,2%	5,8%	7,5%	5,2%	5,0%	5,0%	4.7%	5.0%
Materials Science	2,0%	1,5%	2,0%	1,3%	1,4%	0,5%	0,5%	0,0%	0.0%	0.9%
Mathematics	4,0%	2,0%	0,9%	1,0%	0,0%	0,0%	0,0%	0,0%	0.0%	1.0%
Microbiology	0,9%	2,0%	1,9%	1,9%	0,7%	3,4%	5,3%	3,2%	3.1%	2.5%
Molecular Biology & Genetics	0,5%	0,5%	0,5%	0,5%	0,4%	0,4%	0,5%	1,7%	1.0%	0.6%
Neuroscience & Behavior	0,0%	0,7%	1,2%	1,1%	1,5%	1,5%	1,4%	0,6%	0.4%	1.0%
Pharmacology & Toxicology	4,5%	0,8%	2,9%	3,0%	2,5%	6,4%	6,9%	6,9%	4.6%	4.3%
Physics	1,4%	0,8%	0,0%	0,5%	0,9%	0,5%	0,0%	2,0%	2.3%	1.0%
Plant & Animal Science	3,4%	2,3%	4,3%	3,9%	4,9%	5,6%	4,5%	3,0%	2.7%	3.9%
Psychiatry & Psychology	1,0%	0,9%	0,7%	0,0%	0,0%	0,0%	0,0%	0,5%	0.0%	0.3%
Social Sciences	1,7%	1,2%	1,2%	1,1%	1,0%	1,0%	1,0%	0,4%	0.4%	0.9%
Space Science	1,9%	4,0%	4,9%	7,8%	5,7%	6,4%	3,3%	1,0%	1.1%	4.1%

Table 9 Evolution of the number of HCRs affiliated in France by ESI field and year. The darkest values in the table are the highest. (Color figure online)

French HCRs by discipline and year

	2014	2015	2016	2017	2018	2019	2020	2021	2022	TOTAL
Agricultural Sciences	5	5	7	9	4	4	4	3	0	41
Biology and Biochemistry	2	3	3	2	4	4	2	0	0	20
Chemistry	3	6	7	6	5	4	2	1	1	35
Clinical Medicine	13	11	13	12	21	20	24	26	25	165
Computer Science	2	1	2	2	2	2	2	1	2	16
Cross-Field	0	0	0	0	55	59	60	58	56	288
Economics and Business	2	2	2	3	4	3	3	3	4	26
Engineering	9	2	3	1	0	0	0	1	2	18
Environment and Ecology	9	8	10	6	8	6	8	8	6	69
Geosciences	9	10	11	9	9	8	8	6	5	75
Immunology	0	3	9	8	11	7	10	8	10	66
Materials Science	3	2	3	2	3	1	1	0	0	15
Mathematics	4	2	1	1	0	0	0	0	0	8
Microbiology	1	2	2	2	1	4	7	4	4	27
Molecular Biology and Genetics	1	1	1	1	1	1	1	3	2	12
Neuroscience and Behavior	0	1	2	2	3	3	3	1	1	16
Pharmacology and Toxicology	6	1	4	4	4	9	10	11	7	56
Physics	2	1	0	1	2	1	0	4	4	15
Plant and Animal Science	6	4	9	8	11	11	10	6	5	70
Psychiatry and Psychology	1	1	1	0	0	0	0	1	0	4
Social Sciences	3	2	2	2	2	2	2	1	1	17
Space Science	2	4	5	8	7	7	4	1	1	39
TOTAL	**83**	**72**	**97**	**89**	**157**	**156**	**161**	**147**	**136**	**1098**



Urban area	Cumulated nb of HCR	Part
Paris	571	52.0%
Marseille	90	8.2%
Toulouse	81	7.4%
Grenoble	73	6.6%
Montpellier	50	4.6%
Lyon	49	4.5%
Bordeaux	43	3.9%
Nancy	29	2.6%
Nantes	19	1.7%
Clermont-Ferrand	17	1.5%
Nice	14	1.3%
Strasbourg	11	1.0%
Other urban areas	34	3.1%
Total	**1098**	**100.0%**

Fig. 4 Distribution of HCRs in major university agglomerations. © Tableau Software

over the past 5 years is particularly intriguing. It seems contradictory with the prominent position in this field of several French institutions in the ARWU ranking by subjects and with the number of the Fields medals awarded to their faculty in the last two decades. We discuss this issue in the Discussion section at the end of the paper.

Geographical and institutional HCR landscape

The number of HCRs also enables the analysis of promising institutions and areas at a country level. Bornmann and Bauer (2015) analyzed the 2014 list; they observed that German HCRs concentrate in two Lands (Bavaria and Baden-Württemberg), in six cities, and two institutions (the Max Planck Institute and the University of Heidelberg) and that they are particularly active in the fields of natural sciences, materials science, and chemistry. In China, many HCRs are affiliated with the Academy of Sciences (Wei & He, 2021) and concentrate their research on physics, chemistry, materials science, and engineering (Li, 2018). The concentration of highly cited researchers in a few institutions is an observation widely shared, as observed by Basu (2006), who relies his analysis on an indicator that defines the intensity of HCRs in institutions (IHCR).

As shown in Fig. 4, the French HCR landscape is dominated, as expected, by the region of Paris, home of some of the French flagship universities and outstanding research institutions.

The question of the concentration of research in France is at the heart of the restructuring of the university landscape implemented over the last 20 years, in the framework of the Excellence Initiative program (IDEX). As Bernela and Bouba-Olga (2014) point out, the idea that concentrating efforts on a few territories would make it possible to "benefit from economies of agglomeration and/or reach a critical size, a necessary condition for performance […] motivated the policy of competitiveness clusters, the initial objective of which was to label some fifteen clusters of excellence intended to facilitate interactions between companies, research centers, and training institutions."

However, the benefits obtained from a concentration of research are called into question by the same authors: they consider that "the most accomplished empirical work on the question shows […] that scientific production tends to be deconcentrated, without any significant impact on the quality of the research produced." (Bernela & Bouba-Olga, 2014). In a more recent paper, Grossetti et al. (2020) state that "there seems to be no effect of the

Table 10 Distribution of HCRs France (2014–2022) by Idex (university cluster)

Idex	Cumulated nb of HCR
Outside an Idex	559
AMIDEX (Marseille)	72
BORDEAUX	33
IPS Paris-Saclay	164
LYON	15
PARIS CITE	101
PSL (Paris)	24
SU (Paris)	66
UCA (Nice)	13
UGA (Grenoble)	47
UNISTRA (Strasbourg)	4
Total	1098

Table 11 Distribution of HCRs affiliated in France by year and type of institution. © Tableau Software. The darkest values in the table are the highest. (Color figure online)

	2014	2015	2016	2017	2018	2019	2020	2021	2022
School/University	43	29	42	42	86	139	124	123	114
Research Organi..	24	24	36	32	51	5	16	10	9
Research Institu..	24	2	7	4	6	5	7	2	3
Other	5	7	1	4	6	2	2	2	1
Organization/Fe..	1	2	3	1	3	2	4	6	5
Hospital	9	6	6	7	8	2	3	3	1
Company	1	2	2	3	3	3	5	3	3

geographical concentration of researchers on their 'productivity' measured by the number of publications". They add that researchers in the same city publish similarly whether their institutions are or not clustered (or labeled), and that "scientific output" is a function of the number of researchers and does not depend on any city size effect. In any case, it seems complicated to disentangle the size of a researcher population and the policy drivers (i.e., the time derivative) of relatively high ratios of HCRs to total researchers.

Our analysis of the geographical distribution of the HCRs (Table 10) does not seem to corroborate Grosseti's findings. The fact is that our 'scientific elite' of HCRs is concentrated in the Paris area (more than 50%) while half of the HCRs from the latest annual list belong to one of the ten IDEX described in the list of the French *Observatoire des Sciences et Techniques* (OST, 2021).[2]

Other than the significant increase in the number of HCRs observed in 2018 (introduction of the Cross-Field category), Table 11 also shows a decline in the figures corresponding to the national research organizations from 2019 onwards in favor of academic

[2] This is not exactly the official list since the status of Univ Lyon has not been confirmed yet. One caveat though: the indicator of the number of HCRs is affected here both by disciplinary biases (over-representation of the medical and health sector) and by the imprecision of the location of personnel of national research organizations.

institutions. It is apparent that HCRs working in a joint research unit (UMR) —the general case in France for most researchers— began switching their affiliation from national bodies to academic institutions, arguably to enhance the international visibility of French universities through the Shanghai ranking. According to the official report on scientific employment in France (MESRI, 2023), researchers in national research organizations represent approximately 25% of the workforce in the public sector.

Discussion

The community of highly cited authors undoubtedly constitutes an influential scientific elite whose impact is commensurate to the interest aroused by their publications.

However, many scientific leaders, who are laureates of major scientific prizes or members of the National Academy of Sciences, have not benefited from this recognition: we must thus highlight that the ensemble of HCRs covers only partly the population of the most influential researchers of a given discipline.

The case of mathematics

A striking example, in the case of France, is that of Mathematics: no French HCR was listed in the Mathematics category in recent years, even though several French universities appear at the world Top10 in this field in the subject rankings and very regularly, researchers from its laboratories are winners of the prestigious Fields Medal (7 out of the 22 winners from 2002 to 2022 belonged to French institutions). At a more global scale, the situation is similar: only three universities out of the Top 25 of the ARWU ranking by subject in 2022 have an HCR in the Mathematics category, and prestigious institutions in this specific domain, like Oxford, Cambridge or MIT are conspicuous by their absence of the list for this category.

How to explain this contradiction? We must necessarily return to the process of identification of the HCRs. These researchers must have coauthored a significant number of highly cited papers (HCP) during an 11-year window (2011–2021 for the 2022 list), the threshold being moreover variable from one discipline to another, to adjust to the publication specificities of the corresponding scientific community. We estimate the publication thresholds for Mathematics in the 2022 list to be 11 HCPs, with more than 430 citations.

It is easy to understand that a researcher can obtain strong recognition among peers while having a few highly cited articles and thus be absent from the list. In the same way, a researcher who clears a specialized and innovative field of research, despite being warmly recognized for the accomplishment, may only receive citations from the relatively narrow community of researchers who can understand and follow the new scientific advances and therefore remain below the citation threshold. In other words, the HCR methodology favors the authors of reference articles (or state-of-the-art reviews) aimed at a large (potentially citing) community. The same observation can apply to yearly reports (e.g., on health issues) or articles describing an innovative methodology or software of general interest.

There are also striking examples in other disciplines: the Nobel Prize in Physics 2022 was awarded to Alain Aspect 'for experiments with entangled photons, pioneering quantum information science'. Alain Aspect, a French researcher, is not on the HCR lists, which comes as no surprise since he has made a remarkable breakthrough in a very innovative field in which few international teams have tried their hand and, for this reason, did not attract a large number of citations.

Another aspect of the HCPs is that even if the threshold for each category is adjusted according to the size and publication practices of the discipline (the publication threshold for mathematics being lower than that of the medical sciences, for example), each category itself groups several sub-groups which concern communities of potentially very different sizes: thus an article in applied mathematics which proposes methodological tools will be potentially cited by a much larger community than an article dealing with the fundamentals of the discipline, such as is the case of outstanding contributions in algebra, analysis, geometry, or topology. This issue also arises in other categories. In the social sciences, the determination of HCPs favors the authors of articles dealing with public health because they interest a larger community. In astronomy, papers dealing with cosmology are much more likely to be cited than those dealing with, for example, planetology, simply because the respective publishing and citing communities are of very different sizes (to the benefit of cosmology).

In practice, for France and the publications of the years 2011–2021, the table below, constructed from data from the Web of Science, gives some characteristics of the 'mathematics' category, illustrating that the highly cited articles mainly come from the field of applied mathematics (Web of Science Category: Applied mathematics):

- Total number of articles (Mathematics, 2011–2021): 526,000
- Highly Cited Articles (HCP): 5260
    - Applied Mathematics: 3002
    - Probability, Statistics: 785
    - Fundamental Mathematics: 878
    - Multidisciplinary or Interdisciplinary: 750
    - Computational Biology: 85
- Share of the total showing at least an author with French affiliation: 36,393 articles (i.e. 7% of 5,26,000), and 255 HCPs (i.e. 5% of 5260) classified as follows by subdomain:
    - Applied Mathematics 130
    - Probabilities, statistics: 63
    - Fundamental Mathematics: 54

Indices to assess the proximity of a researcher to the HCR threshold

The question of finding a simple indicator or a composite indicator in order to identify the most prominent researchers in a given disciplinary field has often been asked.

Ioannidis et al (2016) have, for example, proposed a composite indicator that takes into account multiple authorship: they suggest giving greater weight to the single author of an article, or to an author occupying the first or the last place in the list of authors. They also include the Schreiber hm index –which addresses the issue of multiple authorship– into the calculation of the composite index. Although this strategy might seem attractive to identify

Table 12 Case of the discipline 'Geosciences'

HCP/cites rank	Geosciences author	*h*-index	*Q*	*M*	HQM	HQM rank	HCR status
1	Ciais, Philippe	82	25.6	85	244.7	2	Yes
2	Piao, Shilong	36	34.1	93	160.1	11	Yes
3	Friedlingstein, Pierre	41	74.8	212	298.9	1	Yes
4	Berthier, Etienne	43	29.0	86	166.2	9	Yes
5	Poulter, Ben	38	44.4	165	219.8	4	Yes
6	Peng, Shushi	35	29.6	90	148.3	12	Yes
7	Bony, Sandrine	40	28.9	125	179.2	6	Yes
8	Kerr, Yann	43	22.1	53	129.4	14	Yes
9	Bopp, Laurent	47	43.3	134	234.2	3	Yes
10	Viovy, Nicolas	29	35.5	164	169.8	8	Yes
11	Dufresne, Jean-Louis	27	27.1	79	116.2	17	Yes
12	Masson-Delmotte, V	44	32.8	66	161.1	10	Yes
13	Schulz, Michael	41	44.1	111	202.0	5	Yes
14	Colette, Augustin	28	25.7	71	112.6	18	No
15	Chevallier, Frédéric	45	29.1	84	170.4	7	No
16	Al Bitar, Ahmad	31	36.4	75	137.8	13	No
17	Vautard, Robert	33	25.9	57	117.2	16	No
18	Deque, Michel	30	28.8	81	128.3	15	No
19	Donnadieu, Yannick	31	21.1	41	94.1	19	No
20	Claustre, Hervé	29	23.1	38	90.1	20	No

The table presents the top 20 of the list of affiliated researchers in France in this discipline (according to InCites, April 2023), sorted by decreasing number of HCP. Values for h-index, Q, and M (mean number of citations per paper) and their geometric mean HQM are also presented. The last column gives the HCR status: 13 of the Top 20 researchers were listed as HCR in the category Geosciences, with permanent or temporary affiliation in France, from 2014 to 2022

many outstanding authors in a field, it is however not appropriate to reproduce the HCRs of Clarivate's list insofar as their mode of selection neither gives any preference to the order of the authors nor explicitly handles multi-authorship (except for ruling out HCPs with more than 30 authors).

We already mentioned the suggestion of Sinatra et al. (2016) to use a bibliometric measure, the Q-parameter,[3] in order to unlock the secret behind scientific success by untangling the role of productivity, luck, and individual ability in a scientific career. We have tested the ability of the Q-parameter to place HCR authors at the top of the list. However, our tests carried out in several disciplines do not give a particular advantage to the Q-parameter. We present here, by way of example, the counts obtained in the Geosciences discipline for the first 20 authors with the most HCPs and citations in this discipline (according to InCites 2023). Table 12, below, shows the values of the following indices: the Q-parameter,

[3] The *Q*-parameter of the Q-model accounts for the individual-level ability to make high-quality scientific contributions and is not affected by time. The higher the *Q* parameter is, the higher the probability that researchers will publish high-quality scientific papers.

The Q parameter is based on the geometric mean of the citations that a scientific oeuvre of an individual researcher receives (after ten years of publication).

the h-index which rewards consistency in publishing cited papers, and the mean number M of citations per paper.

None of these indicators is enough to account for the ability of HCRs to produce highly cited papers. All three indicators have a different way of rewarding consistency in high citation rates and taking (or not) uncited works into account.

We also calculated a composite indicator, HQM, as the geometric mean of these three indexes (we use $H = h^2$ because it is homogeneous to the number of citations).

A Spearman test based on the four ranks thus determined, shows that the HQM composite presents the best correlation (0.7) with the order determined by the HCPs and their citations, the Q-parameter presenting only a correlation coefficient of 0.5. The potential of the HQM indicator deserves to be tested and discussed on a larger scale than that of the French HCRs alone. This will be the subject of further work.

Note here that while it is relatively simple for a researcher to calculate the indicators mentioned here in order to assess the closeness to the thresholds for inclusion in the list of HCRs, he or she will still have to calculate the scores of those researchers who appear above the threshold to compare against, which requires some familiarity with the intricacies of databases (disambiguation of names among other issues).

Indices based on the number of citations generally suffer from significant limitations, such as self-citation bias and lack of standardization across research fields. The use of indicators within a given field (ESI research area) somewhat limits the issue of standardization, although in an incomplete way, as discussed above in the case of Mathematics.

Conclusion

The analysis of the case of the French HCRs can help us to outline answers to two questions: does the list of HCRs make it possible to identify a scientific elite? Is the number of HCRs a relevant indicator for an institution or a country? To the first question, if the answer is positive, it is accompanied by a reservation: indeed, the list of HCRs incontestably makes it possible to identify a scientific elite of 'super-publishers'; however, it does no satisfactorily help in the search for future winners of major international prizes, or the identification of the scientific leaders of the research front in their field. And indeed, there is little overlap between the HCR list and the lists of holders of scientific awards or national members of the Academy of Sciences.

Among the characteristics of the population thus determined, it should be added that the gender bias against women remains particularly flagrant, that the procedure selects researchers who are most often in the second half of their career (average over 50 years), and finally that communities for which articles in international journals are not at the core of their activity, as well as communities where influential articles currently have more than 30 authors, may have been insufficiently considered. In addition, our study shows that questionable practices (e.g., self-citations, contested and retracted articles) should be carefully scrutinized.

We do not have a clear answer to the second question, HCR as an indicator. On the one hand, the number of HCRs accumulated by an institution reflects the influence of this institution through its overall number of citations and the capacity of this institution to provide its most eminent researchers with the best scientific environment and access to research infrastructures. But we also identified relevant limitations, as in the case of Mathematics, where the indicator seems to contradict the recognition attributed elsewhere (notably through distinctions and medals) to worldwide respected scientific institutions in the field.

Finally, we introduced a new indicator, HQM, calculated as the geometric mean of three previously defined productivity indicators (namely, the square of the h-index, the Q-parameter, and the average number of citations). This composite index could help potential users to assess their proximity to the HCR threshold within their own field of research.

Author contributions LC: conceptualization and data curation, LC, DE, DD: formal analysis and investigation, LC: writing—original draft, DE, DD: writing—review and editing, DE, DD: supervision.

Funding The work of Domingo Docampo has received financial support from the Xunta de Galicia (Centro singular de investigación de Galicia accreditation 2019–2022) and the European Union (European Regional Development Fund—ERDF).

Data availability Web of Science raw data used in this study have been made available under license by Clarivate Analytics. The authors are not allowed to redistribute WoS data which therefore cannot be made available. All other data are available from the corresponding author upon reasonable request.

Declarations

Conflict of interest The authors declare no conflict of interest.

References

Aksnes, D. W., & Aagaard, K. (2021). Lone geniuses or one among many? An explorative study of contemporary highly cited researchers. *Journal of Data and Information Science*, 6*(2), 41–66. https://doi.org/10.2478/jdis-20210019

Alhuthali, S., & Sayed, A. (2022). Saudi universities rapid escalation in academic ranking systems: implications and challenges. *Journal of Controversial Ideas*. https://doi.org/10.35995/jci02010008

Basu, A. (2006). Using ISI's "Highly Cited Researchers" to obtain a country level indicator of citation excellence. *Scientometrics, 68*(3), 361–375. https://doi.org/10.1007/s11192-006-0117-x

Basu, A., Foland, P., Holdridge, G., & Shelton, R. D. (2018). China's rising leadership in science and technology: quantitative and qualitative indicators. *Scientometrics, 117*(1), 249–269. https://doi.org/10.1007/s11192-018-2877-5

Bernela, B., & Bouba-Olga, O. (2014). Concentrer la recherche et attirer des créatifs? Remise en cause d'une politique à partir d'études empiriques. *Annales Des Mines—Responsabilité Et Environnement, 74*, 71–74. https://doi.org/10.3917/re.074.0071

Borchardt, R., & Hartings, M. R. (2018). The academic papers researchers regard as significant are not those that are highly cited. Blog LSE. Retrieved May 14, 2018, from https://blogs.lse.ac.uk/impactofsocialsciences/2018/05/14/the-academic-papers-researchers-regard-as-significant-are-not-those-that-are-highly-cited/.

Bornmann, L. (2014). How are excellent (highly cited) papers defined in bibliometrics? A quantitative analysis of the literature. *Research Evaluation, 23*(2), 166–173. https://doi.org/10.1093/reseval/rvu002

Bornmann, L., & Bauer, J. (2015). Evaluation of the highly-cited researchers' database for a country: Proposals for meaningful analyses on the example of Germany. *Scientometrics, 105*(3), 1997–2003. https://doi.org/10.1007/s11192-015-1619-1

Bornmann, L., & Daniel, H. (2008). What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation, 64*(1), 45–80. https://doi.org/10.1108/00220410810844150

Bornmann, L., & Leydesdorff, L. (2018). Count highly-cited papers instead of papers with *h* citations: Use normalized citation counts and compare "like with like"! *Scientometrics, 115*, 1119–1123. https://doi.org/10.1007/s11192-018-2682-1

Bornmann, L., Bauer, J., & Haunschild, R. (2015). Distribution of women and men among highly cited scientists. *Journal of the Association for Information Science and Technology, 66*(12), 2715–2716. https://doi.org/10.1002/asi.23583

Bornmann, L., Bauer, J., & Schlagberger, E. M. (2017). Characteristics of highly cited researchers 2015 in Germany. *Scientometrics, 111*(1), 543–545. https://doi.org/10.1007/s11192-017-2248-7

Clarivate Analytics. (2022). Highly cited researchers (web site). Retrieved March 1, 2023, from https://clarivate.com/highly-cited-researchers/.

Confraria, H., Blanckenberg, J., & Swart, C. (2018). The characteristics of highly cited researchers in Africa. *Research Evaluation, 27*(3), 222–237. https://doi.org/10.1093/reseval/rvy017

Docampo, D., & Cram, L. (2014). *Highly cited researchers and the Shanghai ranking—technical report*. https://doi.org/10.13140/RG.2.2.28524.16008

Docampo, D., & Cram, L. (2019). Highly cited researchers: A moving target. *Scientometrics, 118*(3), 1011–1025. https://doi.org/10.1007/s11192-018-2993-2

Docampo, D., Egret, D., & Cram, L. (2022). An anatomy of the academic ranking of world universities (Shanghai ranking). *SN Social Sciences, 2*, 146. https://doi.org/10.1007/s43545-022-00443-3

Grossetti, M., Maisonobe, M., Jégou, L., Milard, B., & Cabanac, G. (2020). Spatial organisation of French research from the scholarly publication standpoint (1999–2017): Long-standing dynamics and policy-induced disorder. *EPJ Web of Conferences, 244*, 01005. https://doi.org/10.1051/epjconf/202024401005

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

Ioannidis, J. P., Klavans, R., & Boyack, K. W. (2016). Multiple citation indicators and their composite across scientific disciplines. *PLoS Biology, 14*(7), e1002501. https://doi.org/10.1371/journal.pbio.1002501

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), e3000384. https://doi.org/10.1371/journal.pbio.3000384

Kamali, N., Rahimi, F., & Abadi, A. T. B. (2022). Learning from retracted papers authored by the highly cited iran-affiliated researchers: Revisiting research policies and a key message to clarivate analytics. *Science and Engineering Ethics, 28*, 18. https://doi.org/10.1007/s11948-022-00368-3

Koltun, V., & Hafner, D. (2021). The H-index is no longer an effective correlate of scientific reputation. *PLoS ONE, 16*(6), e0253397. https://doi.org/10.1371/journal.pone.0253397

Kosmulski, M. (2020). Nobel laureates are not hot. *Scientometrics, 123*, 487–495. https://doi.org/10.1007/s11192-020-03378-9

Larivière, V., Ni, C., Gingras, Y., Cronin, B., & Sugimoto, C. R. (2013). Global gender disparities in science. *Nature, 504*(7479), 211–213. https://doi.org/10.1038/504211a

Leydesdorff, L., Wouters, P., & Bornmann, L. (2016). Professional and citizen bibliometrics: Complementarities and ambivalences in the development and use of indicators—a state-of-the-art report. *Scientometrics, 109*(3), 2129–2150. https://doi.org/10.1007/s11192-016-2150-8

Li, J. T. (2018). On the advancement of highly cited research in China: An analysis of the highly cited database. *PLoS ONE, 13*(4), e0196341. https://doi.org/10.1371/journal.pone.0196341

Martinez, M., & Sá, C. (2020). Highly cited in the south: International collaboration and research recognition among Brazil's highly cited researchers. *Journal of Studies in International Education, 24*(1), 39–58. https://doi.org/10.1177/1028315319888890

Meho, L. I. (2020). Highly prestigious international academic awards and their impact on university rankings. *Quantitative Science Studies, 1*(2), 824–848. https://doi.org/10.1162/qss_a_00045

Meho, L. I. (2022). Gender gap among Highly Cited Researchers, 2014–2021. *Quantitative Science Studies, 3*(4), 1003–1023. https://doi.org/10.1162/qss_a_00218

MESRI. (2023). State of scientific employment in France—report 2023. Retrieved from https://www.vie-publique.fr/rapport/288692-l-etat-de-l-emploi-scientifique-en-france-rapport-2023.

Must, Ü. (2020). The highly cited researchers with researcher ID: Patterns of behavior through time. *Journal of Scientometric Research, 9*(2), 195–199. https://doi.org/10.5530/jscires.9.2.23

Nagane, H. S., Fukudome, Y., & Maki, K. (2018). An analysis of star scientists in Japan. In *2018 IEEE international conference on engineering, technology and innovation (ICE/ITMC), Stuttgart, Germany* (pp. 1–5). https://doi.org/10.1109/ICE.2018.8436388.

Nielsen, M. W., & Andersen, J. P. (2021). Global citation inequality is on the rise. *Proceedings of the National Academy of Sciences, 118*(7), e2012208118. https://doi.org/10.1073/pnas.2012208118

OST. (2021). La position scientifique de la France dans le monde et en Europe, 2002–2018. Observatoire des Sciences et Techniques. Retrieved from https://www.hceres.fr/fr/publications/la-position-scientifique-de-la-france-dans-le-monde-et-en-europe-2005-2018-ost.

Rodríguez-Navarro, A., & Brito, R. (2022). Research assessment based on the number of top researchers. *Journal of Scientometric Research, 11*(3), 286–294. https://doi.org/10.5530/jscires.11.3.32

Schreiber, M. (2013). A case study of the arbitrariness of the h-index and the highly-cited-publications indicator. *Journal of Informetrics, 7*(2), 379–387. https://doi.org/10.1016/j.joi.2012.12.006

Shamsi, A. (2021). Gender of Highly Cited Researchers focused on the cross-field category. *Gaceta Sanitaria, 35*(5), 506–507. https://doi.org/10.1016/j.gaceta.2020.03.007

Sinatra, R., Wang, D., Deville, P., Song, C., & Barabási, A. L. (2016). Quantifying the evolution of individual scientific impact. *Science, 354*(6312), aaf5239. https://doi.org/10.1126/science.aaf5239

Sinay, L., Carter, R. W., & de Sinay, M. C. F. (2020). In the race for knowledge, is human capital the most essential element? *Humanities & Social Sciences Communications, 7*, 20. https://doi.org/10.1057/s41599-020-0521-5

Szomszor, M., Pendlebury, D. A., & Adams, J. (2020). How much is too much? The difference between research influence and self-citation excess. *Scientometrics, 123*, 1119–1147. https://doi.org/10.1007/s11192-020-03417-5

Teixeira da Silva, J. A., & Dobránszki, J. (2018). Multiple versions of the *h*-index: Cautionary use for formal academic purposes. *Scientometrics, 115*, 1107–1113. https://doi.org/10.1007/s11192-018-2680-3

Tijssen, R. J. W., Visser, M. S., & van Leeuwen, T. N. (2002). Benchmarking international scientific excellence: Are highly cited research papers an appropriate frame of reference? *Scientometrics, 54*, 381–397. https://doi.org/10.1023/A:1016082432660

Van Leeuwen, T. N., Visser, M. S., Moed, H. F., Nederhof, T. J., & Van Raan, A. F. J. (2003). The Holy Grail of science policy: Exploring and combining bibliometric tools in search of scientific excellence. *Scientometrics, 57*, 257–280. https://doi.org/10.1023/A:1024141819302

Van Noorden, R. (2020). Highly cited researcher banned from journal board for citation abuse. *Nature, 578*, 200–201. https://doi.org/10.1038/d41586-020-00335-7

Van Noorden, R., & Chawla, D. S. (2019). Hundreds of extreme self-citing scientists revealed in new database. *Nature, 572*, 578–579. https://doi.org/10.1038/d41586-019-02479-7

Waltman, L. (2016). A review of the literature on citation impact indicators. *Journal of Informetrics, 10*(2), 365–391. https://doi.org/10.1016/j.joi.2016.02.007

Warner, J. (2000). A critical review of the application of citation studies to the Research Assessment Exercises. *Journal of Information Science, 26*(6), 453–460. https://doi.org/10.1177/016555150002600607

Wei, Y., & He, S. (2021). The characteristics of highly cited researchers in China. In *iConference 2021*. Retrieved from http://hdl.handle.net/2142/109688.

Zhang, N., Wan, S., Wang, P., Zhang, P., & Wu, Q. (2018). A bibliometric analysis of highly cited papers in the field of economics and business based on the essential science indicators database. *Scientometrics, 116*, 1039–1053. https://doi.org/10.1007/s11192-018-2786-7

Zuckerman, H. (1977). Scientific elite. Nobel laureates in the United States. Transaction Publishers. Retrieved from https://wellcomecollection.org/works/wexh9yfx.