

# Working with scientific literature - Impact, quality and finding

Methods of Scientific Working (for Crop Sciences) (3502-440)

21 Jan 2025

## Table of contents

<b>1</b>	<b>Motivation</b>	<b>1</b>
<b>2</b>	<b>Impact and quality of scientific literature</b>	<b>2</b>
2.1	Measuring the impact of science . . . . .	2
2.2	Retractions of scientific publications . . . . .	7
<b>3</b>	<b>Finding scientific literature</b>	<b>8</b>
<b>4</b>	<b>Summary</b>	<b>9</b>
<b>5</b>	<b>Key concepts</b>	<b>9</b>
<b>6</b>	<b>Further reading</b>	<b>10</b>
<b>7</b>	<b>Study questions</b>	<b>10</b>
<b>8</b>	<b>In-class activities</b>	<b>10</b>
8.1	Exercise: Evaluating the Potential of Large-Scale Organic Agriculture . . . . .	10
8.2	Discussion: Problems with the impact factor . . . . .	11
8.3	Doing a Google Scholar search (Demo) . . . . .	12
	<b>References</b>	<b>12</b>

## 1 Motivation

In the first part of our two classes on working with scientific literature we focused on summarizing scientific literature and using outlines and other techniques as well as citing scientific literature correctly in our scientific writings.

In this chapter we focus on working with the literature, meaning that we want to define how we find and correctly use existing literature.

Scientists also frequently evaluate the scientific writing of their colleagues, for example, when reviewing grant applications or conducting peer reviews of scientific publications.

Another central aspect of working with literature is how to assess the quality of existing publications before building our own research projects and scientific writings on it. Which publications are trustworthy, and which tools or measures do we have to evaluate them?

Additionally, when evaluating scientific research for purposes such as developing legal regulations or making policy decisions (e.g., regulations for genetically modified organisms or climate change policies), it is important to understand the criteria by which the quality of scientific literature is assessed.

Closely related to the topic of the quality of literature is the question we can find literature that is relevant for our research topic. Which approaches and tools are available to find the relevant literature and to make sure that we are not missing key publications in our field of research?

## 2 Impact and quality of scientific literature

### 2.1 Measuring the impact of science

Science is an expensive activity and it is increasingly evaluated for its impact on society and economics. Scientists, research institutes and universities are being evaluated for their scientific productivity. For this reason, various measures have been created to measure scientific productivity. They include

- amount of money obtained from external, competitive funds
- the number and quality of scientific publications
- the number of patents

#### The impact of scientific publications

Scientific publications are probably the most important outcome of scientific research because they are submitted to peer review, and only if this review process was successful, a paper appears in a scientific journal. It can then be considered to belong to the scientific literature and the body of human knowledge.

Scientific publications are measured by some metrics that are supposed to reflect their impact. A standard **bibliometric measure** is the **impact factor**, which is defined as the number of citations the articles published in a journal accumulate over a certain period of time. Impact factors can be calculated for scientific journals, but also for research institutes, individual researchers, and even for individual papers.

#### The impact factor (IF)

In its initial definition, the impact factor, often abbreviated IF, is a measure reflecting the average number of citations of articles published in peer-reviewed scientific journals. It is frequently used as a proxy for the relative importance of a journal within its field, with journals with higher impact factors deemed to be more important than those with lower ones. The impact factor was devised by Eugene Garfield, the founder of the Institute for Scientific Information (ISI), now part Clarivate Analytics. Impact factors are calculated yearly for those journals that are indexed in Clarivate's Journal Citation Reports.

In a given year, the impact factor of a journal is the average number of citations to those papers that were published during the two preceding years. For example, the 2008 impact factor of a journal would be calculated as follows:

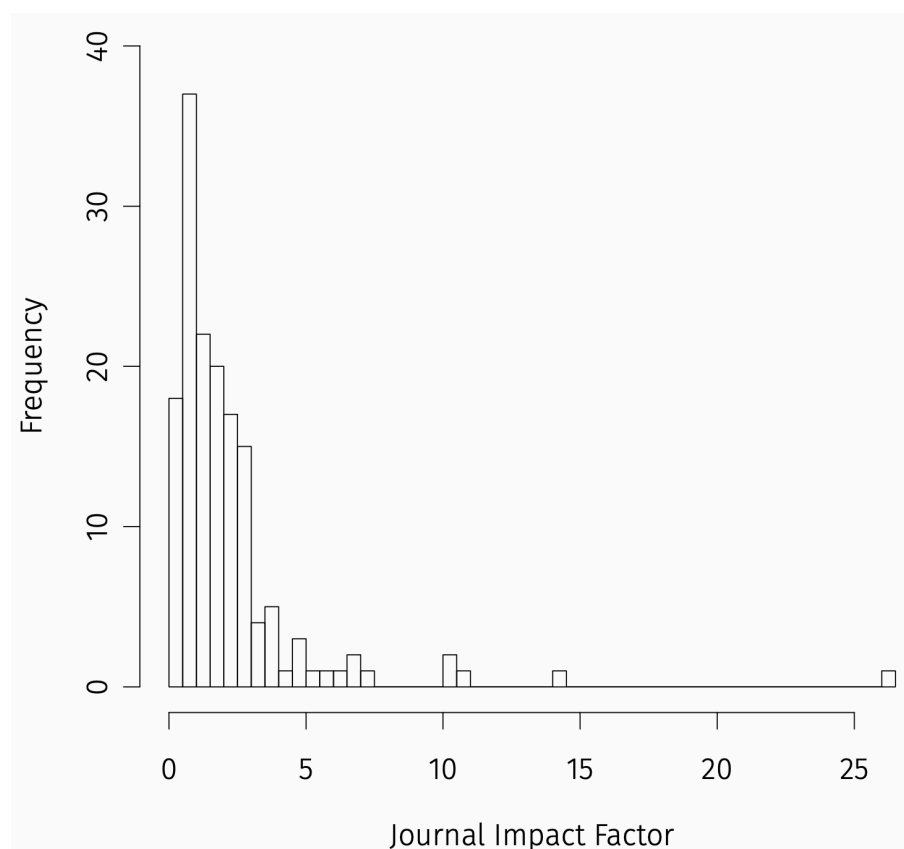
- $A$  = the number of times articles published in 2006 and 2007 were cited by indexed journals during 2008
- $B$  = the total number of 'citable items' published in 2006 and 2007
- The impact factor for 2008 is then  $= A/B$

In summary, the impact factor can be expressed by the following equation:

$$\text{IF} = \frac{\text{Number of citations during year of articles from last two years}}{\text{Number of articles in the last two years}}$$

The impact factors for journals of a particular field can be downloaded from the Journal Citation Reports, which requires a registration ([Link](#)).

Figure 1 shows the distribution impact factors of journals in the field of 'plant science' in the year 2009.



**Figure 1** – Distribution of impact factors for journals in the field of plant science ranked by the height of the impact factor. The impact factors are from the year 2009. Source: Web of Science Website

The distribution is log-normal, which indicates that most journals have a rather low impact, and only very few journals have a high impact factor. The top five journals in the figure are **review journals**: Annual review of Plant Biology, Annual Review of Phytopathology, Current Opinion on Plant Biology, Trends in Plant Science (Figure 1). These journals only publish review papers and no original research. The first journal with mainly original research papers is Plant Cell with an impact factor of 9.88.

Multiple other measures for the impact of scientific research have been developed. They measure different aspects of scientific citations. For example, the **cited half-life index** expresses the median age of the articles that were cited in a journal each year. The higher the half life index the longer (in number of years), the articles of a journal are cited over the years, which can be taken as an indicator of the fundamental importance of publications in such a journal. This is based, however, on the assumption that publications that address fundamental scientific questions become 'classic' papers and are cited longer.

However, most other measures of journal citations are more or less strongly correlated to the impact factor and are not considered further.

It is also possible to analyse the distribution of citations for a given journal, i.e., count the number of citations that an article has received within a given period of time. Importantly, self-citations are also included in this analysis.

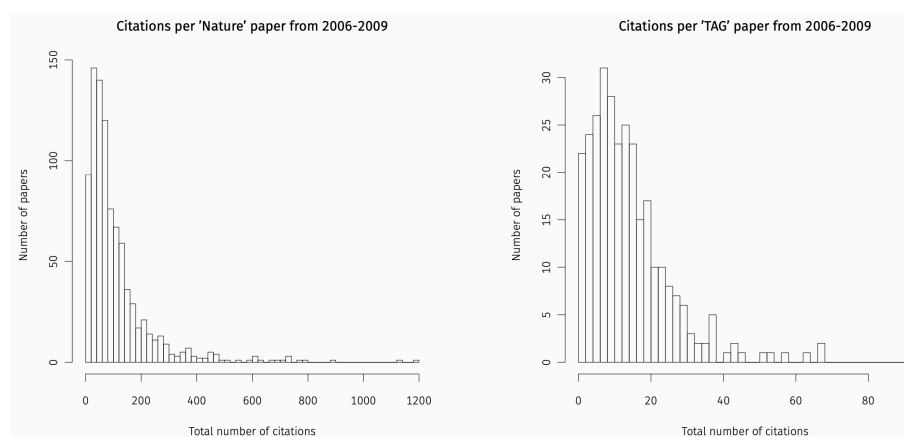
To understand the features of the impact factor, but also its limitations, it is helpful to compare the distribution of impact factors from a journal with a low and a high impact factor. Table 1 shows summary statistics that outline the differences between the two journals.

The number of citations also follow a log-normal distribution (Figure 2).

**Table 1** – Comparison of citations of scientific articles that appeared in a general and a specialist science journal from 2006 to 2009.

Journal name	Impact factor	Articles	Mean	Median	Min.	Max.
Nature	34.480	905	110.6	71	0	1194
Theoretical and Applied Genetics	3.363	298	14.7	12	0	91

The distributions show that high-profile journals derive their high impact factors mainly due to a small number of highly cited papers.



**Figure 2** – Citations per article from 2006 to 2009 in a general (*Nature*) and a specialized (*Theoretical and Applied Genetics*) science journal

### Measuring the Impact of individual researchers

It is also possible to measure and compare the impact of individual scientists, which is now common practice in the hiring process or in the evaluation grant proposal applications. There are several measures to evaluate the impact of individual scientists:

**Total number of peer-reviewed articles published** This measure is easily calculated from records of the author in databases.

**Average impact factor of journals** The impact factor of the journals in which a particular person has published articles. This use is widespread, but controversial because there is a wide variation from article to article within a single journal (Figure 2).

**Total number of citations of a scientist's papers** This can be also easily extracted from curated or automated websites.

**Hirsch-Index (*h* index)** An *h* value of 15 means that an author has published 15 papers which were cited at least 15 times. It was shown that this index may be able to analyse the scientific productivity of an individual at various career stages because expectation values can be formulated.

***b* Index** The *h* index divided by the numbers of years a researcher has spent in science.

Many more statistics have been devised that allow to evaluate scientists without actually having to read their papers.

## The $h$ index explained

The  $h$  index is widely used to measure the scientific output of individual research. It is named after its inventor Hirsch (2005) who developed this measure to enable the ranking of individual scientists.<sup>1</sup>

It has the following definition:

- $n$  articles that are cited at least  $n$ -times
- 1 article 1 times cited:  $h = 1$
- 4 articles 4 times cited:  $h = 4$

### Exercise

Calculate the  $h$ -index of a scientist with the following citation counts:

- 1st article, 19 times cited
- 2nd article, 8 times cited
- 3rd article, 4 times cited
- 4th article, 20 times cited
- 5th article, 10 times cited

## Criticism of impact factors

The main purpose of impact factors is that they allow a formalized, possibly objective and essentially automated assessment of individual scientists, research institutions or scientific journals in the vast and extremely diverse scientific enterprise. Hence it can be considered to be an efficient system of **reputation building**.

However, any such formal system is liable to abuse by adapting to the rules of the game and utilizing (or manipulating) them to maximise one's fitness in this system. For example, the journal impact factor and not the scientific direction of a journal may be used in the decision to which journal a scientific article should be submitted. Numerous criticisms have been made of the use of impact factors. The most basic arguments against using impact factors are

- Quantity of publications does not imply quality
- The real value of research becomes often obvious much later
- Differences in the impact of publications: A single researcher may produce only a single, but outstanding paper in his or her lifetime and not much else, whereas a mediocre researcher may publish many papers with mediocre content.
- Counting impact factors puts a big administrative burden on scientists (although this is done automatically, e.g., with Google Scholar)
- All measures raise the issue of authorship. It is highly advantageous to be a co-author on many papers, even though the contribution to a paper may be minimal. This is because indices do not count the order of authorship or the relative contribution.

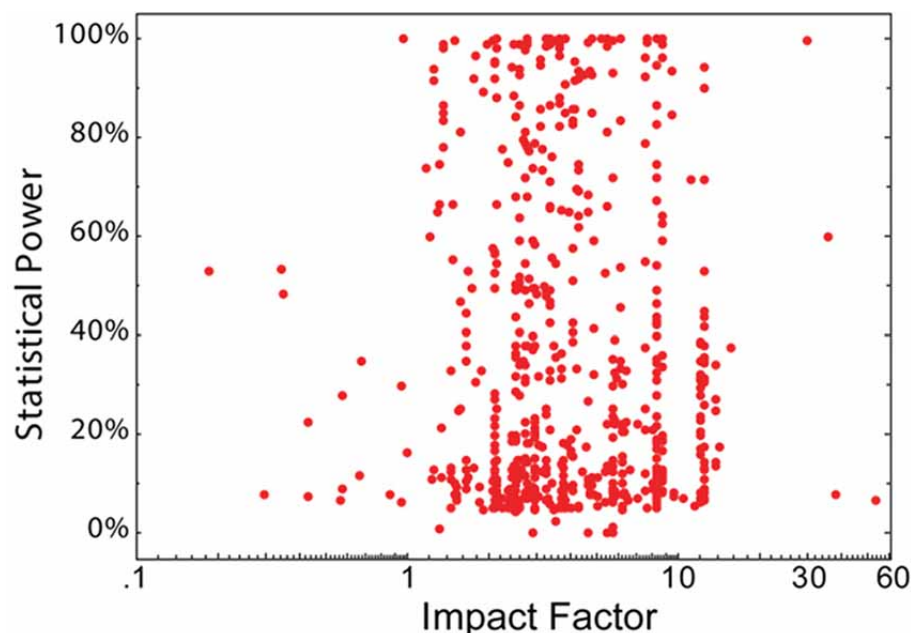
Besides the more general debate on the usefulness of citation metrics, criticisms mainly concern the validity of the impact factor, possible manipulation, and its misuse.

---

<sup>1</sup>See also [Wikipedia](#)

### Low correlation between impact factor and research quality

Quantitative analyses of the statistical power of studies (which is a measure for the quality of the underlying experimental design) and the journal impact factor (IF) in which the research was published revealed a low or no correlation between both measures. This is shown in Figure 3 for the field of neuroscience.



**Figure 3** – Statistical power of 650 neuroscience studies plotted as a function of the impact factor of the journal in which studies were published. Each red dot indicates a single study. Source: Brembs (2018)

Another analysis shows that journals with an above-average error rate rank higher than journals with a lower error-rate.

### A possible alternative to impact factors?

There are many criticisms about using simple metrics to measure the quality of scientists and scientific work. Good historical examples where such metrics would have failed are described in this [blog post](#).

On the other hand, the number of scientific publications is growing very rapidly, and no scientists can follow the complete body of scientific literature within a field of specialty and even less so outside. Therefore, some measure of importance is necessary to evaluate the key papers in a scientific field and the impact factor of a paper and the journal in which it is published is a first step to find out.

To counter the publication of too many scientific papers with too little impact of each, scientific funding agency such as the Deutsche Forschungsgemeinschaft now ask to list only the ten most important publications in a grant proposal.<sup>2</sup> The expectation is that scientist write now longer papers with more complete datasets and analyses, which subsequently may result in a higher impact. In other words, the quality and not the quantity of research is going to become important.

The **Declaration on Research Assessment (DORA)** is an initiative aimed at improving the ways research outputs are evaluated. It emphasizes moving away from traditional metrics like journal impact factors and encourages assessing research based on its quality, impact, and contributions to

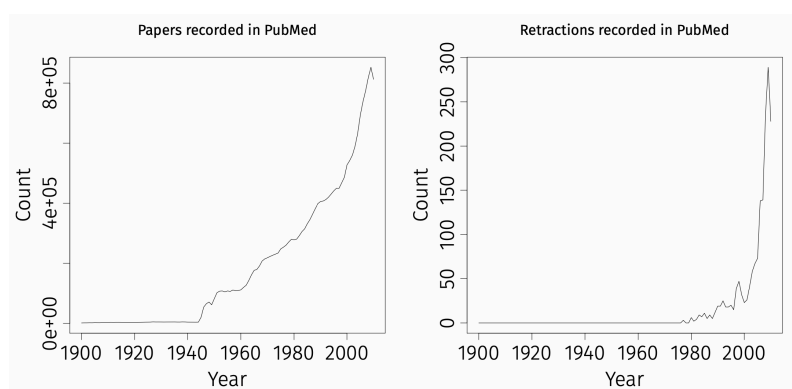
<sup>2</sup>See this [link](#)

the field, regardless of the venue of publication. DORA advocates for more transparent, inclusive, and comprehensive research evaluation practices across disciplines. More information can be found at this [link](#). Many institutions subscribe to this initiative, but it is unclear to what degree this commitment has a substantial impact on the hiring committees at academic institutions.

## 2.2 Retractions of scientific publications

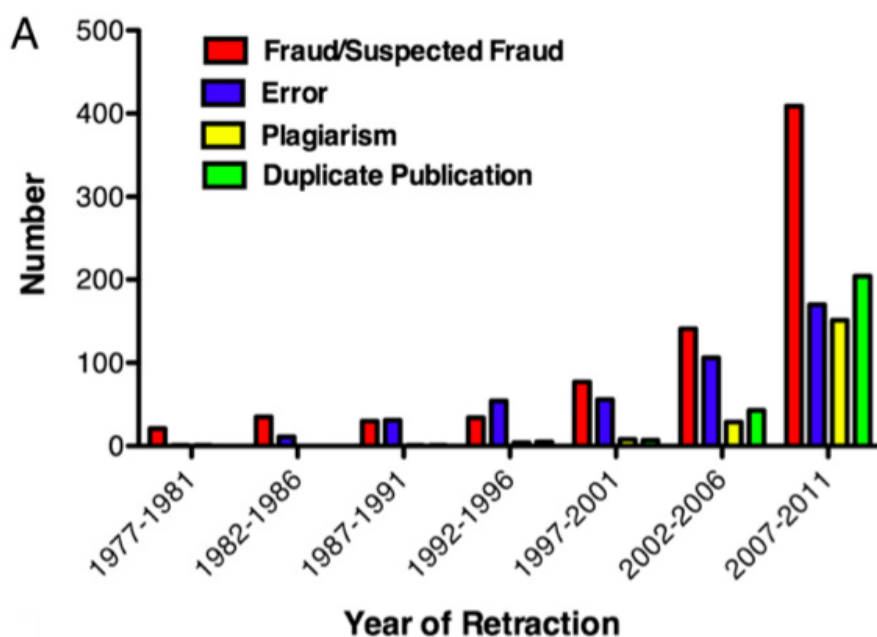
The growing numbers of papers per year is accompanied by an even faster growth of retracted papers (Figure 4). This is due to several reasons:

- The growing awareness for falsifications.
- New techniques to recognize the manipulation of images.
- A stronger competition between research groups that gives a higher chance that critical and important experiments are being repeated.



**Figure 4** – Numbers of papers indexed in the public repository PubMed and numbers of papers retracted

One study investigated the cause of paper retractions (Fang et al., 2012). It shows that a majority of retractions result from fraud or suspected fraud.



**Figure 5** – Causes of retraction of scientific research papers. Source: Fang et al. (2012)

There is also a watch blog that follows up retractions due to scientific fraud or other problems <sup>3</sup>. There seem to be a much higher proportion of retractions in the field of biomedicine than in agricultural research.

It is important to note that wrong, falsified or controversial papers also can accumulate high impact factors. Hence, a high impact factor is not necessarily a sign of a high scientific quality.

### 3 Finding scientific literature

After investigating how to evaluate the quality of scientific literature, it is also important to discuss how to find relevant literature, particularly high-quality scientific literature. Whether you are conducting research to understand the state of the art in your field or writing your thesis or publication, finding the most up-to-date and relevant literature is crucial.

The field of literature search is evolving rapidly. In the past, one approach was to regularly scan the most important journals in the field, often by subscribing to email notifications of new issues, which allowed for a consistent review of interesting and relevant papers. Another approach was to search **curated literature databases** such as Web of Science or Scopus, both of which require a subscription and are owned by major publishing companies. The University of Hohenheim has subscribed to Scopus ([Link](#)), and therefore you can use it for your literature search.

However, the field is changing quickly due to the increasing availability of public information and the growing influence of artificial intelligence. Numerous companies now offer AI-based tools to search for relevant literature, providing criteria to assess the relevance of matches—such as the number of citations a publication has generated or through citation networks that help identify key publications in a field.

In the following sections, we will explore two of these approaches.

The first approach is almost considered a classical method today: using a search engine. **Google Scholar** ([Link](#)) is arguably the largest and most effective search engine for finding relevant papers. Most importantly, it is free to use, and instead of relying on manual curation, it leverages Google's search algorithms to identify relevant publications.

A key feature of Google Scholar is the ability to filter searches by specific years, and once you've found a publication, you can assess its quality by reviewing its citation count—how often it has been cited by others.

Additionally, many scientists maintain a Google Scholar profile where they list their publications. Google automatically calculates an *h-index*<sup>4</sup>, which serves as a useful indicator of a researcher's scientific impact and the quality of their work—in other words, how trustworthy their research is.

ResearchRabbit ([researchrabbit.ai](https://researchrabbit.ai)) searches public databases similar to Google Scholar and analyses citation networks. It has a graphical presentation of these networks that allows to identify publications that are well connected to other publications and therefore of potential importance. It can be linked easily with Zotero and can contribute very well to establishing collections of literature on a given topic.

---

<sup>3</sup>[Retractionwatch](#)

<sup>4</sup>See, for example <https://scholar.google.de/citations?user=TN3bbNwAAAAJ&hl=de>



### AI-based methods for working with the scientific literature.

The field of methods based on searching for scientific literature is rapidly evolving. For this reason, we provide only a few links:

Inciteful can be linked with Zotero and allows to search for the relevant literature on a given topic. See this [blog post](#) for a description and links to other, potentially useful AI-based tools for literature search.

[Semantic Scholar](#) is another tool for literature searches.

#### Advice on literature search for master thesis

In Bachelor's and Master's theses, it is often noticeable that students do not cite the appropriate literature.

For example, instead of citing peer-reviewed research publications or reviews, students cite websites, newspaper articles, or articles from non-peer-reviewed sources. Also, quite frequently, articles from unknown or predatory research journals are cited.

This strongly suggests that the students have not engaged seriously with the subject matter and have mostly compiled their references from regular Google searches—not from Google Scholar searches.

This may lead to a negative assessment of a thesis.

Therefore, an important tip for your Master's thesis: Take the time to search for studies that are either published in the most visible journals or are the most highly cited in the field. Review articles are also acceptable, and you should cite these.

Ideally, you should fully read these articles to grasp their content.

Another important tip is not to simply copy figures from the internet or from articles, such as in the introduction of a Master's thesis. Instead, consider how you can, with minimal effort, create your own diagram or infographic that demonstrates that you have understood, reflected on, and worked through the topic.

## 4 Summary

- While the basic definition of plagiarism is straightforward, the actual classification of the extent of plagiarism is difficult to achieve and requires specialized software tools
- Outlining, i.e., the summarization and the expansion of summarized concepts in one's own word is a good approach to avoid plagiarism of other people's work.
- Bibliometric measures are an automated and standardized way of evaluating research.
- The most widely used measure is the impact factor and various variants of it.
- The impact factor is defined as the number of citations a publication receives per time unit.
- Bibliometric assessment of scientific impact has many problems and is frequently criticized.
- The use of impact factors is based solely on formal criteria; however, the content and therefore creativity and impact is not assessed.
- Wrong or controversial papers can also get high impact factors.
- Alternative measures to evaluate the quality of research publications are being developed that attempt to focus on the content and less on formal criteria for the evaluation.

## 5 Key concepts

- Impact factor

- Hirsch index,  $h$

## 6 Further reading

- Callaway (2016) - A concise summary of the problems with the impact factor.

## 7 Study questions

1. How is plagiarism defined?
2. What different types of plagiarism exist?
3. Which techniques do scientists have available to avoid plagiarism?
4. What is the journal impact factor, and how is it calculated?
5. Assume that your task is to evaluate a controversial study that was published in a scientific journal. Which criteria can you use to evaluate the quality of the researchers and of the publication?
6. Which different measures of impact exist, how are they defined and what are the differences among them?
7. Numbers of citations of journal articles follow an exponential (or log-normal distribution). If the average impact factors of individual scientists, research institutions or journals is compared: What distribution of impact do you expect for those? Why?
8. What possibilities exist to manipulate the impact factors?

## 8 In-class activities

### 8.1 Exercise: Evaluating the Potential of Large-Scale Organic Agriculture

**Scenario:** You are an intern for a member of parliament who is considering a proposal to transition the agricultural system to fully organic, based on polling data indicating public support and the recognized benefits for long-term sustainability.

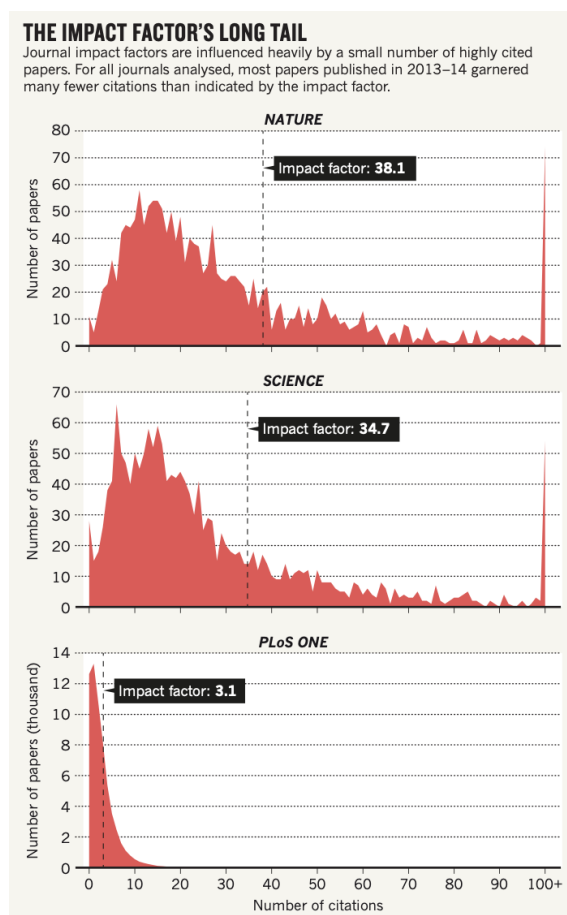
Your task is to assess the scientific evidence to determine whether large-scale organic agriculture can realistically achieve these goals. You are required to write a short, 2-page policy paper, referencing no more than five scientific publications.

Please consider the following questions:

- **Criteria for Selection:** What formal and informal criteria will you use to select the five papers to include in your policy paper?
- **Relevance of Studies:** What will be your key criteria for determining the relevance of a scientific study for this policy paper?
- **Summarization and Referencing:** How will you summarize and reference the selected studies to convey the necessary information effectively?
- **Addressing Unreferenced Studies:** How will you account for the remaining studies that you cannot reference in the policy paper?

## 8.2 Discussion: Problems with the impact factor

The impact factor of a journal in which a scientist publishes is often misused to evaluate the importance of an individual publication or evaluate an individual researcher. This does not work well since a small number of publications are cited much more than the majority - for example, about 90% of Nature's 2004 impact factor was based on only a quarter of its publications, and thus the importance of any one publication will be different from, and in most cases less than, the overall number (Figure 6).



**Figure 6** – Distribution of number of citations in Nature, Science and Plos ONE journals indicating that in the former two, the high impact factor is mainly determined by a small number of publications with a very large number of citations. Source: Callaway (2016)

The impact factor, however, averages over all articles and thus underestimates the citations of the most cited articles while exaggerating the number of citations of the majority of articles.

Discuss the following questions:

1. What is a characteristic of the distribution of the impact factors that all three journals have in common?
2. Why does the shape of the distribution carry only limited information about the average number of citations over two years (i.e., the impact factor) of a random article published in the journal?
3. Which reasons can you think of that may explain the differences in the impact factors of the three journals in Figure 6?
4. Can you summarize the key problems of the impact factor as a measure of scientific quality?

### 8.3 Doing a Google Scholar search (Demo)

## References

- Brembs B. 2018. Prestigious Science Journals Struggle to Reach Even Average Reliability. *Frontiers in Human Neuroscience* **12**. doi:[10.3389/fnhum.2018.00037](https://doi.org/10.3389/fnhum.2018.00037)
- Callaway E. 2016. Publishing elite turns against impact factor. *Nature* **535**:210–211.
- Fang FC, Steen RG, Casadevall A. 2012. Misconduct accounts for the majority of retracted scientific publications. *Proceedings of the National Academy of Sciences* **109**:17028–17033. doi:[10.1073/pnas.1212247109](https://doi.org/10.1073/pnas.1212247109)
- Hirsch JE. 2005. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences* **102**:16569–16572. doi:[10.1073/pnas.0507655102](https://doi.org/10.1073/pnas.0507655102)