UiT
THE ARCTIC
UNIVERSITY
OF NORWAY

Faculty of Humanities, Social Sciences, and Education

# CITATIONS AND RETRACTIONS

*Why are retracted articles cited?*

—

**Julie Hoel**

*DOK-3951 Master thesis in Document Management, May 2016*

ACKNOWLEDGMENTS

I would like to thank Jan Erik Frantsvåg for letting me know about Jon Sudbø and the continued citing of his retracted articles. I would also like to thank my family and friends who gave me words of encouragement when I needed it, even more than I realised myself.

# CITATIONS AND RETRACTIONS

# 1 INTRODUCTION: CITATIONS AND RETRACTIONS

*To stand on the shoulder of giants* is a common phrase, used to showcase the ideal of using other peoples' research to further your own. Citations are a way of showing whose giant shoulders you have been standing on in order to see further into the world of science. This is the ideal for citations – the furthering of scientific ideas by showing not only were you want to go, but also where you came from. Sometimes, however, the giant shoulders you were standing on, is nothing but a pillar of lies. The self-correcting measure of science comes in the form of retractions. When an article has been found erroneous, be it from a simple mistake or more malignant reasons such as fraud and misconduct, said article should be retracted. The whole point of this practice is so that others in the future will not rely upon wrongful work in their research. However sometimes this is exactly what happens.

In 2006 the Norwegian scientist, Doctor Jon Sudbø, considered to be a prodigy in his field, was exposed for committing research fraud. Sudbø was working hard on devising a method to identify markers of oral cancer. His breakthrough came when he made a discovery with the potential of diagnosing oral cancer at a much earlier stage than what was known before. Sudbø gained a lot of attention when he claimed that common painkillers could help prevent oral cancer. Yet they could also be dangerous since they would increase the risk of heart disease. The only problem with his discoveries was that they were simply not true. When people began looking into Sudbø's articles it became clear that his research were largely based upon data which Sudbø had made up. He had gone so far as to create fictional patients with falsified names, social security numbers, and medical histories. When the discovery was made most of his articles were retracted. Thus pushing the giant that is Sudbø's work down, making his shoulders unable to bear a load. However more than one fellow scientist have managed to find perch on the shoulders even after the articles were retracted. Ten years after the scandal Sudbø's retracted articles are still cited.

Jan Erik Frantsvåg, who works as Open Access advisor and researcher at the University of Tromsø – The Arctic University of Norway (UiT), was the one who made me aware of this, and it was his idea for me to take a closer look at the citing of retracted articles, with the case of the Sudbø-fraud as a backdrop. Citation counting, inspired by Eugene Garfield and his promotion of the citation index, has become more and more important. The

value of citation counting and its connection to the impact factor is heavily debated today. It would seem that a high number of citations are the same as, if not quality, at least impact. This puts pressure on scientists perpetuating the idea that you have to publish or perish. Some people have used this pressure as justification for their misconduct. Sudbø himself claimed that: *"he lost his critical thinking in his strive for academic recognition*" (Jakobsen, 17.01.2014)

It is all about getting your research out to the public. The pressure to publish is evident, seeing as everything from recognition, to tenure, and funding can depend upon it. One way to achieve more citations seems to be not only to publish something groundbreaking or controversial, but also to publish in a journal with a high impact factor. However being published in a high ranking journal does not necessarily equate to quality. In fact as the impact factor goes up, so does the retraction rate. (Ferric C. Fang & Casadevall, 2011) If the idea of retractions is to remove it as part of the scientific body, the question must be asked: why are retracted articles cited?

## 1.1 WHAT WE KNOW SO FAR

In order to better understand the citing of retracted articles I began by conducting a search of *Google Scholar* and *Web of Science* to find articles about retraction. This was to better understand the current state of retracted articles, and how they are cited. Before I started working with the subject I knew next to nothing about it. I guess in my naïve mind the idea of retractions had not occurred to me. It is so easy to think that scientific articles that are published are well-researched and without errors. However this would imply that every researcher and scientist are infallible people with a great sense of ethics and morals. Yet some of them cheat, lie and falsify data, seemingly without any form of moral quandary. When this type of fraud is discovered ideally their articles and scientific findings are retracted. Another reason for retraction can be simple errors and mistakes. Scientist, like the rest of us, are not infallible, they make honest mistakes too. Mistakes which can lead to retractions. In fact retractions are not unheard of in scientific publishing, but perhaps my naiveté stems from the fact that you do not hear about it that much. Others have written about the citing of retracted articles. Often by seeing how a certain article is cited after retraction. One of the common

factors seems to be the agreement that not nearly as many articles that should be retracted are, and even if the retraction rate has increased tenfold in the past years (F. C. Fang, Steen, & Casadevall, 2012), the retraction number should be higher. This might not be so unreasonable since only 1 out of every 10 000 articles that are published are retracted (van der Vet & Nijveen, 2016). Another common thread is the agreement that retractions are a difficult thing to handle. Specially the citing of retracted articles, unless you know that a specific article is retracted, there is no easy way of finding it out. Many calls for a tool or program that can trace retractions, thus you avoid the citing of them. The emphasis on following the COPE-guidelines in order to make people aware of retractions is also an area of import. These are subject areas I will look at closer as I explore the citations of retracted articles. I will use Sudbø's retracted articles to gain an understanding of how people cite retracted article. Areas of import include citation indexing, and also the citers' motives. The hope is to better understand the process of citing, why we cite, and why citations have become so important in the recent years through the emphasis of the impact factor. By gaining insight in these areas I hope to come a step closer into understanding not only how, and why we cite, but mainly why retracted articles are cited.

## 1.2 THE LAYOUT OF THE PAPER

I will begin by doing a literary review of retractions, and the citations of retracted articles in chapter two. Here I will also define citation indexes and how the importance of citations counting have become more prominent in the last years as more weight is put on the impact factor. In order to understand why Sudbø is cited I will also define different motives for citing in chapter two. The context of this examination is the case of Jon Sudbø, therefore I will outline his research career and eventual downfall in chapter four. I will do this to show the extent of his misconduct to show the reason for why his articles were retracted. As part of the examination of Sudbø I will do a citation count of two of Sudbø's retracted articles. In conclusion I will discuss my findings of the citation count and try to determine why and how Sudbø's retracted articles are cited in chapter five.

# 2 RETRACTED ARTICLES AND CITATION COUNTING

The retraction of articles are not limited to one scholarly discipline, they occur in all fields and all affiliation countries (Grieneisen & Zhang, 2012). Yet the extent of published articles are not necessarily that big if you compare it to the overall amount of published articles. Paul E. van der Vet and Harm Nijveen estimate that about one in 10 000 published articles are retracted. There is still a worry that continued citing of the articles that are retracted could propagate the retracted results further (van der Vet & Nijveen, 2016). Svend Bruhns write about retractions in his book *Citationsindexering*, and how research articles containing errors are not always retracted. Sometimes because the errors of an article is not confirmed, and sometimes an article is not officially retracted even though it contains errors (Bruhns, 2010, p. 353). Every so often articles that are officially retracted are cited Sometimes as if a retraction notice have not been issued at all (Fulton et al., 2015). Scientific journals which have published articles that are going to be retracted, should publish a notion letting people know about the retraction. Databases who catalogue scientific articles and have registered the original article should also take action. Since retracted articles are no longer part of the body of science, it is important to make the retractions known in order to prevent the spreading of retracted results (van der Vet & Nijveen, 2016). One way to do this is to follow the COPE-guidelines for retraction (COPE, 2009), if they are a member of COPE, which the journals with the highest impact factor generally are. This is to make sure that as many people as possible know about the retraction, to prevent future citations of the retracted article. When Balhara and Mishra checked retracted articles on mental disorder and the effect of the COPE guidelines on retraction, they discovered that "*a significantly higher proportion of articles cited at least once post-retraction were without a freely accessible retraction notice*" One key issue was that the retraction notice was freely accessible, since significantly fewer retracted articles with the retraction notice freely

> *"Even after retraction notice has been issued, authors and peer reviewers still, presumably, unknowingly, cite the article as though no retraction notice had been issued."*
>
> (Fulton, Coates, Williams,

accessible were cited (Balhara & Mishra, 2014). When doing a study to determine reasons for retraction, Steen discovered that 31.8% of the retracted articles in the study were not in compliance with the COPE-guidelines, since they did not have any mention of the retraction notice in the PDF. This is usually done with some kind of watermark, which these did not have (Steen, 2011). Another issue is that not all authors use a digital copy of an article when citing. Sometimes a researcher has filed away a hard-copy of an article for "future research". If they are diligent and try to check to see if an article has been retracted, they do not have a guaranty that a simple search in some database will show that. van der Vet and Nijveen found that in their experience, none of the popular search engines like *Google Scholar*, *Scopus* or *Web of Science* added a warning to a retracted paper in their list of search results. Others have found that they did, but that the adding of a retraction warning was not done consistently (van der Vet & Nijveen, 2016). In 2010 Steen found that retractions were mainly due to error. Out of 742 English language research papers retracted from the PubMed database between 2000 and 2010, 73.5% of the retractions were due to error, or an undisclosed reason. 26.6% of retractions were due to fraud. Yet he noted that the incidences of research fraud was increasing (Steen, 2011). In 2012 he and Ferric C Fang, and Arturo Casadevall came to the same conclusion: that fraud was increasing. They found that scientific misconduct accounts for the majority of retracted scientific publications (F. C. Fang et al., 2012). They conducted a review of all retraced biomedical and life-science research articles indexed in PubMed in May 2012 and discovered that 21.3% of retractions were due to error. 43.4% were due to fraud, 14.2% were duplicate publications and 9.8% was plagiarism. Miscellaneous reasons or unknown reasons accounted for the rest. Fang et al. blames incomplete, uninformative, or misleading retraction announcements for a previous underestimation of the role fraud plays in retractions. Since 1975 the percentage of articles retracted for research fraud have increased nearly tenfold. It seemed as if a small number of authors were responsible for multiple retractions, where authors who had more than ten retractions, almost always where retracted because of fraud. In terms of the citation of retracted articles they found that some retracted articles continue to be cited as valid, whilst other experience a rapid and sustained decline in citations after retraction. They noted that there were articles shrouded in considerable doubt, where there was a high suspicion of misconduct, yet the articles had not been retracted. The

number of articles who are retracted because of fraud, could therefore potentially be higher. The rate could also be higher, since many journals does not include the reason for retraction, therefore it is difficult to determine if the retraction is due to misconduct. (F. C. Fang et al., 2012)

They also point out the correlation between a high journal impact factor and a high retraction rate. A point which Fang and Casadevall revisits in the editorial "Retracted Science and the Retraction Index". If you look at all of the scientific articles that are published today, retractions are a relatively rare occurrence. When Fang and Casadevall looked at the over 28,000 articles over the course of 40 years, *Infection and Immunity* had issued only 15 retractions. Of these retractions eight were retracted because of misconduct, and six were retracted by the authors because of errors, while the remaining articles were retracted due to extensive plagiarism. Even though they point out that retractions are a rare occurrence, they also note that manuscript retractions seem to be happening more often these days. This could be due to an increase in misconduct, but it could also be the result of "*increased detection due to enhanced vigilance*" (Ferric C. Fang & Casadevall, 2011, p. 3855). Fang and Casadevall also point out that: "*most scientist feel that research misconduct is uncommon*", yet a meta-analysis of survey data found that 2% of scientist reported they have committed serious research misconduct at least once, and one-third admitted that they had engaged in questionable research practices. Again prompting Fang et al. to question if the retractions we know of today, are only scratching the surface of the subject. If an author or an article is suspected of misconduct it is not the journals job to investigate, the primary responsibility of investigation is on the institution of the author.

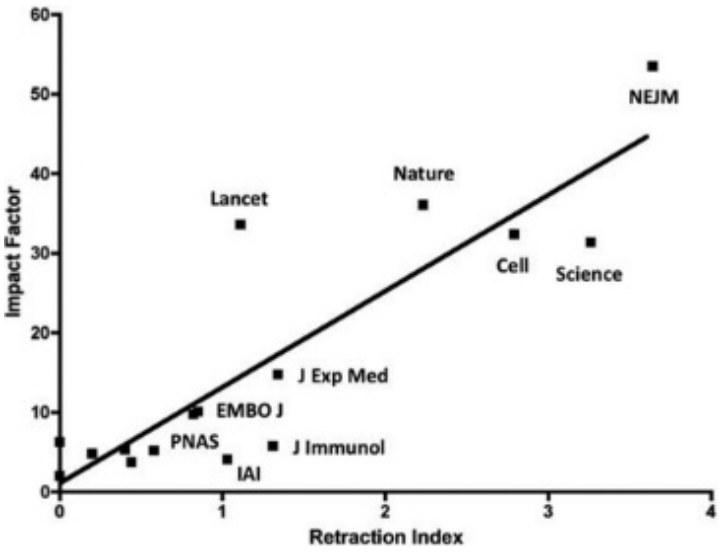



**Figure 1**

CORRELATION BETWEEN IMPACT FACTOR AND RETRACTION INDEX[1]

[1] (Ferric C. Fang & Casadevall, 2011, p. 3856)

However an editor needs to stay vigilant and report if any suspicion of misconduct should arise. When an institution has conducted their investigation it is up to the editor and the journal to react accordingly. Both publishers and authors can request a retraction, and a retracted article should report who has issued the retraction, why, and it should be clear that an article is retracted by following the COPE-guidelines. To investigate if there was a correlation between retraction frequency and the journal impact factor, they searched PubMed for retracted articles among 17 journals with an impact factor from 2.00 to 53.484. They devised a retraction index for each journal by taking the number of retractions from 2001 to 2010, multiplied by 1000, and divided by the number of published articles with abstracts. What they found was that there was a clear correlation between the journal retraction index and the journal impact factor. A reason for this correlation might be the economic and social benefits associated with publishing in high impact journals. Today publication in high impact journals can lead to improved job opportunities, increased likelihood of receiving grants, peer recognition, honorific rewards, even if it is a widespread notion that the impact factor is a flawed measurement of scientific quality and significance. This means that the benefits of publishing in high impact journals can make scientist more prone to commit misconduct. Another reason may be the demand and desire from the high impact journals for clear and definite results, which may encourage authors to manipulate their data in order to present a clear-cut result, when the reality of everyday science may not yield such straightforward conclusions. The other side of the coin is of course the fact that higher impact journals may have increased visibility so they may be under greater scrutiny, so that misconduct is easier to spot (Ferric C. Fang & Casadevall, 2011). Murat Cokol et al. also mentions the higher level of scrutiny that high impact journals go through, and questions if more articles would be retracted if they underwent the same scrutiny (Cokol, Iossifov, Rodriguez-Esteban, & Rzhetsky, 2007). Claiming, like so many others that not all articles that should be retracted are retracted (Budd, Sievert, Schultz, & Scoville, 1999).

> *"…this preliminary investigation suggests that the probability that an article published in a higher-impact journal will be retracted is higher than that for an article published in a lower impact journal."*
>
> (Ferric C. Fang & Casadevall, 2011, p. 3856)

Whatever the reason for the correlation, Fang et al. suggest further studies of the subject, also claiming that the correlation between impact factor and retraction rate, is yet another reason against using bibliometrics as a measure for scientific performance. In conclusion Fang and Casadevall are troubled by the increase in retracted scientific articles. Erroneous and fraudulent research has major ramifications for scientist who are lead astray by falsified results, an unfair distribution of scientific resources, and the potential harm it could cause due to inappropriate medical treatments of patients. Not to mention the chink in the armour of the public confidence in science. The scientific process relies on trust, and retractions and the misconduct of scientist erodes away at that trust. Yet, the infallible scientist does not exist, and retractions are science way of correcting mistakes (Ferric C. Fang & Casadevall, 2011). Even so, this self-correcting measure does not always work. In 2015 Helmar Borneman-Cimenti, Istvan S. Szilagyi and Andreas Sandner-Kielsing released their article "Perpetuation of Retracted Publications Using the Example of the Scott S. Reuben Case: Incidences, Reasons and Possible Improvements", where they investigated the continued citations of Scott S. Reuben, after his work had been retracted due to fraud. They differentiated between citations that noted the retractions and the ones that did not. Reuben's work was retracted in 2009, but have been cited 274 times in a five year period after retraction. Bornemann-Cimenti et al. found that in only 25.8% of the citing articles was it made clear that Reuben's work had been retracted. There was a sharp decline in the number of citations in 25% of the journals, as well as a decline in citations within Reuben's field after the retractions in 2009. Yet Bornemann-Cimenti et al. found that even five years after retraction, nearly half of Reuben's articles are still cited, with only one quarter of the citations mentioning the retraction. Retracted articles should no longer be considered part of the current literature, therefore the continued citations of retracted articles for the people who build upon, since it can lead scientist astray and result in unwarranted recommendations. One reason for the citation of retracted articles Bornemann-Ciment et al. attribute to the fact that people may be unaware of the retractions, especially for authors outside Reuben's field of perioperative pain management. Their data showed that in 2009 72% of citations came from authors with an anaesthesia or pain association, whereas in the years after the percentage varied between 36% and 60%. Another explanation is again connected to the COPE-guidelines, meaning that some

citations after retraction stems from not being aware of the retraction. However in their study they found that only one letter, or 4% of the total number of articles included, was not labelled as retracted. Yet the availability of multiple online versions of the same article, can be the cause of citations, since not all of these make a note of the retraction, leaving readers unaware of the papers status. Another reason for the citation of retracted articles can be the same as van der Vet and Nijveen bring up (van der Vet & Nijveen, 2016). That some authors collect and file articles in order to quote them later, thus not being aware of changes, or retractions that might have happened after they originally found the literature. Finally Bornemann-Cimenti et al. bring up the reviewing process, and the expectation put upon them to catch mistakes, such as the citation of a retracted article. However they point out that the reviewers and editors may be experienced researcher, yet they are not immune to making the same mistakes, or in this case, not catching the same mistakes that authors were unable to catch. Automated software that help in the research and reviewing process might be a solution to this, by checking for note of retraction, but as Bornemann-Cimenti et al. notes, such tools are not yet available. Until such tool are available the only solution is to make people more aware of the situation and the problem of continued citing of retracted articles (Bornemann-Cimenti et al., 2015). Van der Vet and Nijveen also believe that computer programs may be of help in lessening the citing of retracted articles. This requires publishers to structure the retraction status in such a way that it will be readable to a computer programme. Before that time comes, vigilance is required, and one must keep tracking retractions. Blogs such as Retraction Watch aids in this tracking, since automated methods are not yet available (van der Vet & Nijveen, 2016).

> *"In conclusion, the analysis of Reuben's publications clearly shows that retracted articles are still cited 5 years after."*
>
> (Bornemann-Cimenti, Szilagyi, & Sandner-Kiesling, 2015)

## 2.1 THE COMMITTEE ON PUBLICATION ETHICS (COPE)

When people write about retractions one of the most common thing is for them to emphasize the importance of letting people know about the retraction. One ideal method is to follow the COPE and The International Medical Journal Editor guidelines.

In 1997 a small group of medical journal editors in the UK established The Committee on Publication Ethics (COPE). The goal was to provide advice to editors and publishers on all aspects of publication ethics. One particular important feature was to provide a framework on how to handle research and publication misconduct. COPE has over 10 000 members' worldwide, members who are all expected to follow the Code of Conduct for Journal Editors. The goal of COPE is to promote transparency in scholarly publishing. One of the services they provide are guidelines to journals editors on how to handle retractions (COPE, 2016).

### 2.1.1 RETRACTION GUIDELINES

The COPE retraction guidelines instruct the journal editors in when and how to retract. The purpose of retraction is to have a mechanism for correcting the literature and alerting readers about publications with such grave errors that they are not reliable, and to avoid the duplication of data. The main goal of retractions is to correct the literature in order to ensure its integrity. According to the guidelines journals should consider retraction of a publication if they have evidence of errors either from misconduct, like in the case of Sudbø, or with data fabrication, or honest mistakes such as a miscalculation or experimental error. A retraction should also happen if the findings have been published elsewhere without proper cross-referencing, permission or justification, or due to plagiarism or unethical research. Journal editors can also issue expression of concern and corrections. When something has been retracted the journal must provide proper notification of this. According to the guidelines this means that a notice of retraction should:

- *be linked to the retracted article wherever possible (i.e. in all electronic versions)*
- *clearly identify the retracted article (e.g. by including the title and authors in the retraction heading)*

- *be clearly identified as a retraction (i.e. distinct from other types of correction or comment)*
- *be published promptly to minimize harmful effects from misleading publications*
- *be freely available to all readers (i.e. not behind access barriers or available only to subscribers)*
- *state who is retracting the article*
- *state the reason(s) for retraction (to distinguish misconduct from honest error)*
- *avoid statements that are potentially defamatory or libellous (COPE, 2009)*

## 2.2 CITATION COUNTING

Citations are a mandatory part of scientific publishing. It provides us with a link between documents. The scientist does his research and build upon what was there before him, using already established techniques and concepts to further the progress of his field. This dependency of what came before is acknowledged through bibliographic references as "a silent code of honours" (Bellis, 2009, p. xvii) Doctor Eugene Garfield is considered a pioneer in the field of bibliometrics, the statistical analysis of a written publication, such as an article or a book. Citation analysis is an integral part of bibliometrics. Citation analysis in its simplest form consist of counting how many times something or someone have been cited by others. These citation counts can be used to demonstrate the position of an author in his field. It can be used to see who is citing who, and thus determine if there are any links between different documents. For instance if one group of authors are citing each other multiple times, one could assume that there is a strong link between the citing documents and the cited ones. Citation analysis really blossomed in the 1960s mostly due to Garfield and his citation indexes (Bruhns, 2010, pp. 17-19). The ideal for Garfield, which he proposed in 1955, was a bibliographic system that could eliminate the uncritical citation of wrongful data, be it fraudulent, incomplete or obsolete. To do this, Garfield felt that the criticism of earlier papers must be made available through citation indexing (Garfield, 2006a, p. 1123). The hope was to give researchers a way to evaluate the impact of their work, speed up the research process, spot scientific trends, and trace the history of modern scientific thought. In 1958 Garfield laid

the foundations for the Institute for Scientific Information (ISI), now known as Thomson ISI, a dominant force in terms of citation indexing. Scientific indexes at that time were discipline-oriented, which meant that researchers struggled to find the information relevant to their work, since relevant papers from other disciplines were ignored when researching a scientific area by only using subject or keywords. Searching through published scientific papers for relevant information was therefore a labour-intensive process, which seldom yielded the desired results (Yancey, 2005). *The Science Citation Index* was first available in print form in 1964, and two years later the *Science Citation Index* became available on magnetic tape. Later it was released on CD-ROM, along with the *Social Sciences* and *Arts & Humanities* citation indexes, before it emerged as we know it today – the *Web of Science*, the core database in *ISI Web of Knowledge.* The tool which Garfield had envisioned in 1955, was finally realized, and today citation indexing allows researchers across the world to further their studies and evaluate their work (Yancey, 2005). In today's scientific publishing you can search in citations, quickly determining who is citing, and where citations have been cited before (Bruhns, 2010, p. 14). Citation analysis is quite common today. One of the reasons is the idea that receiving a lot of citations is seen as a sign of quality (Bruhns, 2010, p. 17).

## 2.3 IMPACT FACTOR

Indexing scholarly work by citations allowed researchers to track what other works a paper was referencing. As well as keep a count of how many times other people had cited a paper. By counting these citations you had a way of measuring the *impact factor*. The impact factor became an indicator of quality to more influential works. (Yancey, 2005)

The Journal Impact Factor was created by Irving H. Sher and Eugene Garfield. The idea was to give smaller journals a way to stand out against the bigger ones. So instead of measuring the publication count, they would measure the citation count. This allowed for the ranking of journals and authors based on their citations, in order to determine their impact. The journal impact factor is calculated by the number of times all items published in the journal in a two year period is cited by indexed publications, divided by the total number of citable items by the journal in the same two-year period. The two-year period, which the calculations are based upon, is there to make the measure as current as possible. Garfield argues that the ideal

use for the journal impact factor is for researchers who can use it to discover where to find the current reading list of their field, or for information analysis and bibliometrics by tracking bibliometric and citation trends and patterns. It is also there for librarians to help them decide which journals to purchase, or for authors trying to decide where to publish their articles. In other words higher impact might be seen as more prestigious. (Garfield, 2006b) The level of prestige associated with a higher journal impact factor is not without its critics. Many feel that too much weight is put on the impact factor.

> *"As a general rule, the journals with high impact factors include the most prestigious. Some would equate prestige with high impact."*
>
> (Garfield, 2006b, p. 92)

## 2.4 MOTIVE FOR CITING

In 1978 Janet Beavin Bavelas at the University of Victoria, was concerned with the reasons for citing and the growing emphasis on citation counting. She viewed citations and the citation indexing as a way to trace an idea forward, however with some limitations. One must keep in mind that a much-cited article will produce a lot of articles to read in order to trace the idea. Some of these citation sources will be irrelevant or trivial citations. The result is limited to the extent of the journals indexed. Clerical errors such as misspelled names listed as a separate author or article, or duplications will affect the citation count as well. These problems are relevant not only when tracing the underlying idea behind a citation, but also when using citations to measure scholarly impact. Bavelas claims that: "*When citation counts are used in the service of the "sociology of knowledge" or to make decisions about individuals or departments, broader social issues are raised and require consideration at another level.*" (Bavelas, 1978, p. 159) Bavelas questions the objectivity of this practice. This way of indexing citations may be objective, but the underlying initial reasons for citing may not be. "*Only if this process is an objective one, free of subjectivity or social consensus, can data derived from it possibly have the same, desirable characteristic."* (Bavelas, 1978, p. 160). When citing there is a plausibility that there is some level of personal bias or social pressure. At the extreme end of this bias could be the true scholarly impact, where there is a

significant use of the cited authors theory, method or paradigm, but also the attempt to inflate others impact by for example citing the journals editors work, or boosting friends and colleagues. In between these two reasons Bavelas lay the idea of citing to *show your knowledge of the field, and demonstrate your understanding of the fields' most important work.* This form of citing will be guided by a general, not always well-established consensus of what constitutes the important work in the field. It is expected that you know the most pertinent work in your field. Yet what is defined as the most pertinent work in the field tends to be what *everyone* in the field *agrees* is the most pertinent work, and may not always be connected to scholarly impact. Bavelas notes that this is often the case when writing, and both she and others make a large proportion of the citations in order to demonstrate an understanding of the field. (Bavelas, 1978) It is like Bellis describes the function of citations not only to call back to already established ideas and techniques, but also to demonstrate an intellectual link to the ones who came before you. The act of citing becomes a kind of currency where you borrow others credibility and use it to support and reinforce the view that you are advocating. The ability to cite others to strengthen your own claims, helps strengthen your credibility as an expert in the field (Bellis, 2009, pp. xvii-xix). If your goal is to simply strengthen your own claims by citing the sources you are expected to cite, there can sometimes be doubt if the article being cited was even read. In other word citations are not always the same as scholarly impact, because we cite for a variety of reasons. Using citation counts as a measure for scholarly impact only works if the citing author actually read and learned something that is important to bring forward in the article they are writing (Bavelas, 1978). In "Stochastic modelling of citation slips" by Simkin and Roychowdhury they conclude that when many scientist write a manuscript they pick a selection of three papers, cite them and then simply copy a quarter of their references at random, meaning that a large number of citations can be a result of a stochastic process. They came to this conclusion by looking at the distribution of misprints and their ranking by frequency of their repetition. In the end they concluded that as many as 70-90% of scientific citations are copied from the reference list used in other papers, without having been read (Simkin & Roychowdhury, 2005). Lie and Rousseau suggest that this might be pushing the stochastic methods too far. Simkin and Roychowdhury suggest that only 20% of the citers read the original, but copying

a reference is not the same as not reading the article (Liu & Rousseau, 2013). In terms of articles that are cited because of social consensus, one might suggest that articles that are cited, even if it is only because it is expected, does have scholarly impact because everyone agrees that the work is important in the field. This kind of vague and subjective social consensus is the very thing Bavelas is arguing against, since the original purpose of citation counting was to avoid this. The point Bavelas was making was that citations is not an objective indicator of scholarly impact as long as citations are made for other reasons (Bavelas, 1978). More people than Bavelas have been concerned with the motivation for citing, particularly citing in order to increase the visibility of your own work or your colleges. At its base the motive for citing are usually based upon developing, supporting, or applying a concept or a method. Motive for citing can also be to refute an article, claiming that its findings were wrong, but also to review an article as a relevant part of the literature. The other side of citing are the unethical reasons. Not just the social consensus citing that Bavelas talks about, but the conscious cross-citing of others, in a sort of *if you cite me – I will cite you* kind of way. This can make it difficult for new authors to get any recognition, since it makes it hard for them to be cited. The role of the Matthew effect plays a part in terms of citations. (Liu & Rousseau, 2013) It would seem as if those who already are highly cited are more likely to be cited again, making it harder for the lesser known scientist to get recognition. Even in the case of co-authors, the lead-author may not receive as much attention if one of the co-authors are more famous. People will often remember the name of the person that they have heard about before, over an unknown lead-author. (Merton, 1968). More citations means a higher impact factor. The idea of counting citations to measure impact, or how influential something is, is also, according to Bruhns not always correct since citations is not necessarily the same as impact. One reason for this is that citations are counted through scientific citations indexes, when in reality there exists other forms of citations. Scientific indexes such as Thomson Reuters *Web of Science* only indexes scientific articles. They exclude newspaper article and so on. (Bruhns, 2010, pp. 21, 24) The main criticism against the impact factor is the fact that is so limited, yet so widely used today to measure import. Today the journal impact factor is used as concreate measure for scientific quality. Many journals use their impact factor as ranking measure, advertising themselves with how high their impact factor

is. Just because a journal has a high impact factor it does not automatically mean a high citation count for a single article published in that journal. The least cited articles are judged to be as important as the higher cited article in journals with a high journal impact factor. Which goes against the original purpose of the journal impact factor – to separate the more influential works from the lesser ones (Bruhns, 2010, pp. 270-276). The emphasis on the impact factor has led some to cite with impure motives. In 2013 Thomson Reuters suspended a number of Brazilian journals from their journal ranking. The journals in question had tried to artificially boost their own impact factor by publishing articles containing hundreds of references to papers in each other's journals. By not citing articles in their own journals, but instead collaborating with other editors to cite articles in other journals the hope was to avoid detection from analysis that detect extreme self-citation. However they were discovered by Thomson Reuters who had developed a program that specifically target this type of concentrated bursts of citations from one journal to another, also known as *citation stacking*. Four out of 14 Brazilian journals had their impact factors suspended by Thomson Reuters. Mauricio Rocha-e-Silva, who got fired from his position as editor in chief from the journal *Clinics*, claimed that the citation stacking was a result of a growing frustration with Brazils increasing fixation on the impact factor. An agency in the Brazilian education ministry, called CAPES, use the impact factor of the journals that students publish in, to evaluate graduate programs. According to Rocha-e-Silva this means that most graduates will choose higher ranking journals to publish in, over emerging Brazilian journals that are in the lowest ranks. This in turn prevent local journals from improving, since it encourages local researchers to publish in high-impact journals. Rocha-e-Silva claims that this frustration from the editors tempts them to artificially boost their own journals impact factor. In the end Rocha-e-Silva acknowledge that what he did was wrong, but also that the journal impact factor scores are a bad way to judge a scientists work, and CAPES policy of doing so, is the real problem (Noorden, 2013). When commenting on the discovery of the *citation stacking* of the Brazilian journals, Marie McVeigh, the leader of Thomson Reuters annual report for journal citation patterns and impact factors, said that they cannot measure the motive or intent behind a citation, only the patterns and data of a citation: "*Until we can algorithmically measure motive, we are just going to measure citations*" (Noorden, 2013, p. 511).

In fact some of the debate about citation analysis and its worth stems from the lack of knowledge about citers motives. Underlying reasons for citing can be political background, personal motives, or selected referencing in order to promote careers, departments, friends, bosses etc. This goes both ways in that one can under-cite without actually stealing results, but also over-cite in order to inflate the value of a college etc. (Brooks, 1985, pp. 223-225). In 1984 Terrence A. Brooks analysed current competing theoretical models used to describe citers' motivations, coming up with seven key-reasons for citing:

1. *Currency scale*: the citing to show how up-to-date the author is in their field. By reviewing what is most current, the author hopes that his own work is regarded as an advancement of his field.
2. *Negative credit*: authors cite in order to criticize, correct, disclaim or dispute others work. Negative citing is disputed since some believe it does not exist, because erroneous material will be ignored and thus fades into obscurity. On the other hand work of little value, will not be criticized, but a highly cited paper may contain a fruitful error that makes a major contribution to the field.
3. *Operational information*: authors cite in order to reference a concept or a theory, borrowing mathematical or physical techniques, references, results or conclusions from the cited paper.
4. *Persuasiveness:* the author cite in order to support his own view, to persuade others and establish a knowledge claim.
5. *Positive credit*: the citations pays homage, support claims, confirms data and gives credit.
6. *Reader alert*: letting readers know about new, obscure sources, in order to provide background reading and identify original publications.
7. *Social consensus*: introduced by Bavelas, as the social expectations and drive that makes us cite certain sources, based upon vague consensus in a field of study. (Brooks, 1985, p. 226)

When doing a survey to categorize which of these reasons where most likely to be the drive behind citing, what stood out to Brooks was how much *persuasiveness* was a motive for citing. It ranked high both among the humanities and the science fields. In other word a strong

motive for citing is to substantiate your own views, meaning that authors are likely to scour the literature for evidence in support of their own idea. The reason for citing which were the least common ones where the social consensus and the negative credit one. To Brooks this might be because many academic disciplines are loosely organized, therefor lack a strong social consensus, whilst the negative credit motive was almost completely ignored as a reason for citing. For many authors *reader alert* and *operational information* seemed like obscure motives for citing. *Currency* and *positive credit* might be stronger motives because of their possibility to act as surrogates for *pervasiveness*. In the end Brooks concludes that the citer motives are complex and not always something we can put neatly in a box. (Brooks, 1985, pp. 226-228)

# 3 METHOD: THE SUDBØ-CASE AND CITATIONS

In his book *Casestudier: forskningsstrategi, generalisering og forklaring*, Svein S. Andersen defines case studies as "intensive qualitative studies of one or a few investigative unites." Case comes from the Latin word "casus", and it underlines the importance of a single instance – or a single case. In other word a case study generally focus on one action, one procedure, one statement or one turn of events etc. A case study is not based upon statistical procedures, but instead focuses on analytical reasoning combined with additional knowledge about empiricism and theory. (Andersen, 2013, p. 14)

Case studies have grown in popularity since the 1980 because of the possibilities they give. Through as case study you gain context specific knowledge which can be typical and representative of a phenomenon (Andersen, 2013, pp. 23-26). The context, or case that I have examined is Jon Sudbø and his misconduct. I have outlined his background, mainly focusing on the events of 2005 and 2006, when he was under investigation for falsifying data in his research. Case studies are ideal for drawing conclusions in light of an analytical context, which allows you to generate theories in terms of developing terminology, hypothesis and casual mechanisms relevant to the social sciences. (Andersen, 2013, p. 26) By studying the Sudbø-case, and specifically doing a citation analysis of his retracted article, I hoped to say something about the phenomenon that is retracted articles, ultimately having developed a hypothesis to why retracted articles are cited. The citation analysis focused on how Sudbø was cited after retraction. The goal was to determine if there was a specific way of citing him that might justify the citations after retraction, or if the citations came from not knowing that Sudbø's articles are retracted.

## 3.1 CITATION ANALYSIS

When doing a citation analysis Bruhns highlights its advantages by the fact that citation analysis are unobtrusive, in the sense that the people or the units that you are researching are not aware of this. Citation analysis are objective as well, though Bruhns emphasise that this is only true if you use a citation index such as ISI *Web of Science*, since they only index journals that are ISI certified. Citation analysis does however come with some

limitations. Literary references does not show the complete scientific process. Citation analysis are therefore not ideal for tracing the cognitive influences of a specific scientific work. (Bruhns, 2010, pp. 265, 362) Citation analysis as a method is debated because it emphasise the idea that influence is the same as quality. In modern science the value of scientific work is based on how essential it is for further work. (Bellis, 2009, pp. 244, 336) Yet most agree that bibliometrics and citation analysis is not ideal for the evaluation of individual scientists or departments, at least not as the only criteria for evaluation. The use of citation counting does for example not provide you with any knowledge about work that is uncited. Too much emphasis on citation can at its extreme cause a change in citing behaviour as citers motives may change. (Bruhns, 2010, p. 363) Citations counting can however show patterns in citation that are ideal in finding out why retracted articles are cited.

## 3.2 CITATION COUNT

The website *Retraction Watch* is a blog which keeps track of the retraction of scientific articles. The purpose of the website is to make people more aware of the retractions in scientific publishing. One of the ways of doing this is by keeping a list of highly cited retracted articles. What is notable about many of these retracted articles is that they are still cited, even after retraction (Oransky, 2015). The most cited article "Visfatin: A protein secreted by visceral fat that mimics the effects of insulin" (Fukuhara et al., 2005) have, based on numbers from *Web of Science*, 1279 citations, where 776 of them are after retraction. Jon Sudbø and his articles are not cited as much, but he has still received citations even after retraction. The most cited article of Sudbø is "DNA Content as a prognostic marker in patients with oral leucoplakia" (Sudbø et al., 2001). It has 188 citations according to *Web of Science*. Of these 188 citations 40 of them came after retraction.

The citations after retraction could be attributed to people commenting on Sudbø committing research fraud. The Sudbø-case received a lot of publicity, so many people cite Sudbø in the context of talking a about research misconduct. On the other hand some of the citation use Sudbø as source, without acknowledging the retraction. How common this is I have determined by studying the frequency and nature of the citations of Sudbø's work after retraction. Are the citations based upon critiquing Sudbø, or do they mention the fact that

Sudbø's work has been retracted? Or do they simply use him as a legitimate scientific source? By comparing Sudbø's most cited article published in *New England Journal of Medicine* in 2001 (Sudbø et al., 2001) and retracted in 2006, to *the Lancet* article (Sudbo et al., 2005) that finally exposed Sudbø, I will see if there is a difference in the way that these two articles are cited. Both the *New England Journal of Medicine* (NEJM) and *the Lancet* are COPE members. This means that they should adhere to the COPE retraction-guidelines when retracting. The aim of the analysis is to determine how Sudbø is cited after retraction and how different databases adhere to retraction guidelines, in order to exemplify the effect of retractions on the literature.

I used *Web of Science* and *Google Scholar* to determine which of Sudbø's articles was the most cited one. "DNA Content as a Prognostic Marker in Patients with Oral Leucoplakia" was cited 188 times in *Web of Science*, and 327 times according to *Google Scholar*. *The Lancet* article "Non-steroidal anti-inflammatory drugs and the risk of oral cancer: a nested case-control study" is cited 49 times in *Web of Science* and 102 times in *Google Scholar*. I began by determining the yearly citations of the two articles through *Web of Science*s *Citation Report* feature, to see how the two articles were cited both before and after retraction.
I used both *Web of Science* and *Google Scholar* to count the citations, because *Web of Science* only indexes scientific publications, while *Google Scholar* indexes more broadly. This means that the citation count from *Google Scholar* is higher, but also from less reliable sources. Afterwards I compared the differences between the *Web of Science* and the *Google Scholar* results.

I chose to define the search from 2007 to 2016, in the hopes of excluding articles who cited Sudbø before the retraction. However, there were some articles that fell into what I would define as a *grey area*. These were articles that were submitted for review after *the Lancet* incident. These articles were submitted, and sometimes accepted and published in the time between February 2006 and November 2006. The Sudbø-scandal began in January of 2006, when *the Lancet* article finally was publicly exposed. On this basis one could assume that some of the people involved in the publishing process, be it the authors, reviewers or the editors, should have been aware of the Sudbø-investigation, and perhaps taken the time to include an acknowledgment in their papers. However, since the "DNA Content as a Prognostic Marker in Patients with Oral Leucoplakia" was not officially retracted until

November 2nd of 2006, when every co-author except for Sudbø's brother, Asle Sudbø, and Sudbø himself called for a retraction, I have given them the benefit of the doubt and placed them in the *grey area* column. Some articles that cite Sudbø I did not have access to due to language barriers, since I am limited to English and the Scandinavian languages, Norwegian, Danish and Swedish. A very limited number of articles were inaccessible to me because it was content which I would have to pay for, meaning that I could only check the articles available to me via to the access granted to me as a student at the University of Tromsø – The Arctic University of Norway (UiT).

I classified the citations into the following categories:

1) **Debate**: Citations who cite in order to talk about the fraud or the Sudbø-case: these include a lot of ethical debates about scientific misconduct, who discuss the Sudbø-case and how it happened.
2) **Does not mention retraction**: Citations who use Sudbø as a source without acknowledging retraction.
3) **Mentions retraction**: Citations who reference Sudbø, but acknowledge retraction, these were often literary reviews concerned with the current status of the field.
4) **Grey area:** articles submitted and published after the investigation of Sudbø began, that does not acknowledge the ongoing investigation of Sudbø.
5) **No access**: Citations which I did not have access to because of language barrier or because they cost money.

Finally I checked different databases to see how they had handled the two articles in terms of the retractions, and if they had adhered to the recommendations of the COPE guidelines for retraction. I chose to check *Web of Science*, *Google Scholar*, *PubMed*, *Scopus, ProQuest* and *Medline* since these are some of the most commonly used databases in terms of scientific research. What I looked for was if the retraction was noted in the article search record, or noted in the search record title, whether the database linked to the retracted version of the articles, and if the linked to version (PDF) of the articles showed that it was retracted. The final search through the citations were conducted in week 16, 2016, while the checking of adherence to retraction guidelines was finished in week 18, 2016.

# 4 JON SUDBØ: RISE AND FALL

Jon Sudbø graduated at the top of his class as an oncologist (cand.odont) in 1985 and as a doctor (cand.med) in 1994. In 1993, Doctor Albert Reith recruited him to conduct a study determined to find methods to predict oral cancer. Sudbø was given a three-year research grant from the Norwegian Cancer Society to work on his doctoral project, with Reith as his main supervisor. The project took place from 1993 to 2001. (Ekbom et al., 2006, p. 45) During this time Sudbø along with a number of co-authors wrote and published several articles in major and prestigious scientific journals such as *New England Journal of Medicine*, *The Journal of Clinical* and *Oncology*.

On the 15th of October 2005, *The Lancet* published the article "Non-steroidal anti-inflammatory drugs and the risk of oral cancer: a nested case-control study" written by Jon Sudbø as the main author along with thirteen co-authors (Sudbo et al., 2005). The findings of the article did not go unnoticed since they claimed that the long-term use of non-steroidal anti-inflammatory drugs (NSAIDs) would reduce the chance of oral cancer (Sudbo et al., 2005). NSAIDs are drugs who offers pain relief, reduces fever and inflammation, such as ibuprofen, naproxen and aspirin. However the article also claimed that the use of these common painkillers

OCTOBER 2005: *The Lancet* article is published. The article claims that common painkillers can prevent oral cancer, but increases the risk of heart disease.

DECEMBER 2005: Camilla Stoltenberg reads and begins to question the article, raising her concerns to others.

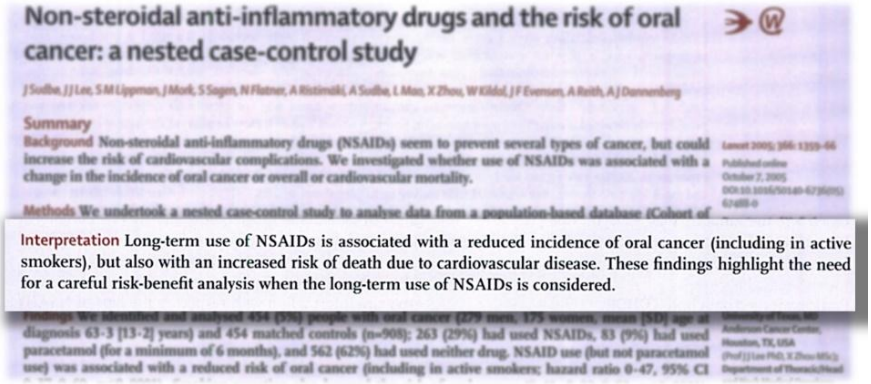
JANUARY 2006: Lars Vatten sends an e-mail to Jon Sudbø, where he questions the findings of the article.

10TH OF JANUARY: Meeting with Sudbø, his main supervisor Albrecht Reith and the Norwegian Cancer Registry.

12TH JANUARY: Sudbø admits to committing research fraud.

JUNE 2006: The report from the research committee is released, concluding that most of Sudbø's work is based on falsified data. Sudbø loses his job, his right to do research, and his medical

would increase the risk of heart disease. (Tunstad, 2011, p. 185) The public reacted accordingly and people became afraid to use these painkillers, and reports kept coming in about patients changing to different, harder drugs, because of fear of using NSAIDs (NTB). The findings of the article did not only bring worry and upset to the public, fellow scientist were beginning to raise questions of their own. One of these people were Division Director with the Norwegian National Health Institute, Camilla Stoltenberg. In December of 2005 she was reading *the Lancet* article and could not for the life of her understand how Sudbø had reached some of his conclusions, when the material he was basing his findings on did not exist. Sudbø had used Norwegian health databases in his study. From the database CONOR (Cohort of Norway) he had collected 454 cancer patients, who were registered as heavy smokers. As a control group he used 454 noncancerous heavy smokers from the same database. The problem was that Stoltenberg was legally responsible for the data in the CONOR study. She knew that Sudbø did not have access to the data, since the data was yet to be released. Not only that, the way Sudbø used and describe the CONOR data was completely wrong. Sudbø claimed to have access to data about the non-steroidal anti-inflammatory painkillers from the Norwegian Prescription Database, data whom Stoltenberg also noticed Sudbø was describing wrong. Errors Sudbø made included saying that the CONOR data went all the way back to 1975, when in reality they were exclusively collected in the 1990s. The data from the Norwegian Prescription Database, Sudbø claimed dated back to the 1970s, when in actuality the Prescription Database was not created until 2004 – *after* Sudbø's collection of data for the


Non-steroidal anti-inflammatory drugs and the risk of oral cancer: a nested case-control study

J Sudbø, J J Lee, S M Lippman, J Mork, S Sagen, N Flatner, A Ristimäki, A Sudbø, L Mao, X Zhou, W Kildal, J F Evensen, A Reith, A J Dannenberg

Summary

Background Non-steroidal anti-inflammatory drugs (NSAIDs) seem to prevent several types of cancer, but could increase the risk of cardiovascular complications. We investigated whether use of NSAIDs was associated with a change in the incidence of oral cancer or overall or cardiovascular mortality.

Methods We undertook a nested case-control study to analyse data from a population-based database (Cohort of

Interpretation Long-term use of NSAIDs is associated with a reduced incidence of oral cancer (including in active smokers), but also with an increased risk of death due to cardiovascular disease. These findings highlight the need for a careful risk-benefit analysis when the long-term use of NSAIDs is considered.

diagnosis 63·3 [13·2] years) and 454 matched controls (n=908); 263 (29%) had used NSAIDs, 83 (9%) had used paracetamol (for a minimum of 6 months), and 562 (62%) had used neither drug. NSAID use (but not paracetamol use) was associated with a reduced risk of oral cancer (including in active smokers; hazard ratio 0·47, 95% CI

**Figure 2**

SUDBØS DOWNFALL: *THE LANCET* ARTICLE WHICH FINALLY EXPOSED JON SUDBØ[2]

[2] (Sudbo et al., 2005)

article. (Tunstad, 2011, pp. 188-189) Stoltenberg conferred with fellow scientist at the Norwegian Institute of Public Health. The Cancer Registry of Norway, which Sudbø had listed as the supplier of the material concerning the cancer cases, was also contacted. They in turn contacted Professor Lars Vatten MD at The Norwegian University of Science and Technology (NTNU), who was a member of the management group for CONOR. (Ekbom et al., 2006, p. 43) Upon reading the article Vatten reacted in disbelief, acknowledging that as groundbreaking as Sudbø's finding may be – none of it was true! Yet Sudbø was given the benefit of the doubt. Surely he had simply made a mistake? A huge one at that, to Stoltenberg this seemed more likely than the idea that Sudbø intentionally had fabricated data and committed fraud. Vatten decided that the best way to gain some clarity in the situation was to go directly to the source. On the 5th of January 2006 he sent an e-mail to Jon Sudbø asking to know more about where the materials of his study was coming from. Upon further investigation Vatten discovers that the Norwegian Data Protection Authority had not given any permission to Sudbø or his colleagues to use data from the different registries Sudbø claimed to have had access to. The Norwegian Cancer Registry had never even received an application to deliver the data Sudbø claimed to have. It is was soon becoming more and more clear that it was not a case of a naïve scientist who had made an error. No it was a case of dishonesty and fraud. On the 10th of January, Vatten, along with colleagues, meet with Jon Sudbø and Albrecht Reith at the Norwegian Cancer Registry. They questioned Sudbø, and they agreed that Sudbø should give up a disc containing all the relevant information to the case. Sudbø seemed tired and left halfway through the meeting. On January 12th Sudbø comes clean and admits that the data used in *the Lancet* article is not true. (Tunstad, 2011, pp. 190-192) On the 18th of January 2006 the Investigation Commission in charge of finding out the extent of Sudbø's fraud was appointed by Rikshospitalet – Radiumhospitalet MC and the University of Oslo.

> *"On the 12th of January, Jon Sudbø admitted to his superiors at the Norwegian Radium Hospital that he had fabricated the computer file which the science results presented in the Lancet article was based upon. Which means that the alleged patients which the analyzed data originates from are fictitious"*
>
> (Ekbom, Helgesen, Tverdal, Lunde, & Vollset, 2006, p. 43)

## 4.1 THE SUDBØ-COMMISSION

Sudbø's employers, Rikshospitalet – Radiumhospitalet HF and the University of Oslo, appointed an investigation commission to determine whether Sudbø had committed research-fraud or not. The committee was established on the 18th of January 2006, releasing its findings on the 30th of June 2006. After admitting to the Investigation Commission that the patients in *the Lancet* article were fictitious, Sudbø also confessed that there were deficiencies with two of his other articles:

- Sudbø J, Samuelsson R, Risberg B, Heistein S, Nyhus C, Samuelsson M, Puntervold R, Sigstad E, Davidson B, Reith A, Berner A. Risk markers of oral cancer in clinically normal mucosa as an aid in smoking cessation counseling. J Clin Oncol. 2005 Mar 20;23:1927-33
- Sudbø J, Lippman SM, Lee JJ, Mao L, Kildal W, Sudbø A, Sagen S, Bryne M, El-Naggar A, Risberg B, Evensen JF, Reith A. The influence of resection and aneuploidy on mortality in oral leukoplakia. N Engl J Med. 2004 Apr 1;350:1405-13.

In light of this, and the hubbub of *the Lancet* article, the Investigation Commission decided it was necessary to investigate all of Jon Sudbø's scientific production and his published work. The Commission set forth to investigate the role of Sudbø's co-authors and contributors as well. The goal was to find the answer to some key-questions:

- *Were the patients who allegedly have been studied real or fictitious?*
- *Was patient data manipulated?*
- *Did serious methodological flaws exist?*
- *Was there evident and serious flaws in the research reporting?*
- *Were there other serious breaches of scientific practice/good research practice? (Ekbom et al., 2006, p. 44)*

The Commission began by mapping the data which was the basis for Sudbø's PhD degree project. This was because the material collected for the PhD had been the basis for Sudbø's subsequent research (Ekbom et al., 2006, p. 44).

In total the Commission went through Sudbø's entire scientific body of work from 1993 to 2006. They examined at least 38 publications and 60 co-authors. The co-authors were all

cooperative, submitting voluntary written statements. The Commission gathered information from all relevant institutions and partners and compared every available data lists and published research results to judge whether the underlying data, which Sudbø's published work was based upon, was real. The Cancer Registry of Norway, which Sudbø cited as a source for his data collecting, also conducted their own internal investigation. Both the Sudbø-Commission and the internal investigation of the Cancer Registry came to the same conclusion:

*"the data underlying parts of Sudbø's PhD project, as well as several other publications, are not sufficiently consistent with the actual facts the Commission has found it reasonable to take into account.*" (Ekbom et al., 2006, p. 6)

In conclusion the Commission determined that the errors they found were too grave to be anything other than fabricated or manipulated to fit the desired findings of Sudbø. The end result was that the doctoral dissertation as well as all publications based on the same raw material, which was most of them, was retracted. In short the bulk of Jon Sudbø's publications was based upon fabrication and manipulation, and therefore invalid. Jon Sudbø was given the chance to read and comment on the findings of the Commission, but failed to provide any information that would change the conclusion of the Commission. They concluded that Sudbø acted alone in his misconduct, since he maintained full and sole control of the underlying data. This level of independence was criticized by the Commission, and they pointed out that the lack of supervision from the Norwegian Radium Hospital was one of the reasons to why Sudbø was able to get away with his misconduct for so long. The co-authors had little to no knowledge about the underlying data material, and most communication that took place went through Sudbø. Therefore the co-authors had little opportunity, or reason, to check the data, or to check who contributed with what. However the Commission pointed out that there

> *«…the Commission has noted that colleagues, researchers, clinicians and individual patients have probably used Sudbø's research results, and it is therefore reasonable to assume that some of them have been affected. The serious implications of this must have been obvious to Jon Sudbø right from the start."*
>
> (Ekbom et al., 2006, p. 8)

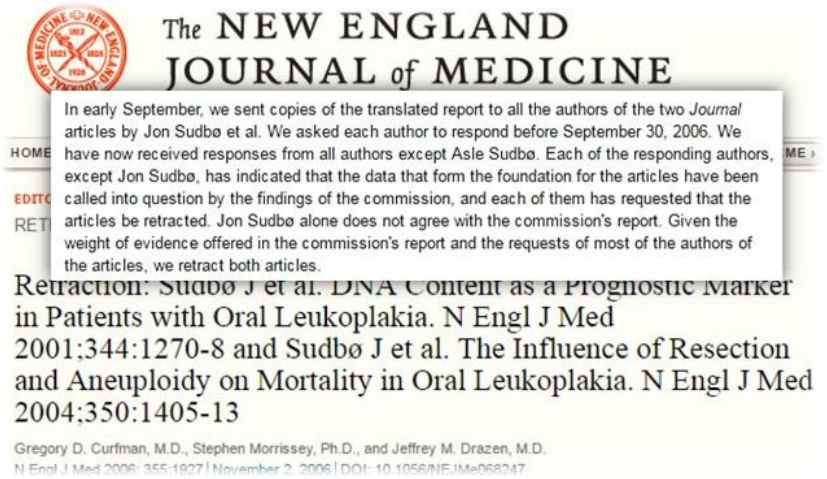
were plenty of signs which people should have picked up on, be it the co-authors, supervisors, superiors, opponents, colleagues and others, ultimately calling it a systemic failure, holding the institutions responsible. They recommend that the institutions in the future take more responsibility for instructing their researchers in proper scientific conduct, as well as exercising some level of verification and control in the research process. The Commission did not view it as their task to assess the ramifications of Sudbø's fraud, but they did point out that colleagues, researchers, clinicians and patients had probably used Sudbø's research results, meaning that at least some of them had been affected (Ekbom et al., 2006, pp. 5-8).

## 4.2 THE MISCONDUCT

The Sudbø-Commission concluded that all of Jon Sudbø's published articles where invalid, except for four original articles, mainly concerning lesser reviews and letters of less scientific value, as well as works in which Jon Sudbø was only a co-author, i.e. publications which mainly have been prepared by others (Ekbom et al., 2006, p. 91).

One of the articles that underwent serious scrutiny from the Commission was, Sudbø's most cited article to date: "DNA Content as a Prognostic Marker in Patients with Oral Leukoplakia", published in the prestigious *New England Journal of Medicine* (NEJM) (Sudbø et al., 2001).

On the 9th of February New England Journal of Medicine (NEJM) also published an expression of concern, following the investigation of Sudbø.


The NEW ENGLAND JOURNAL of MEDICINE

In early September, we sent copies of the translated report to all the authors of the two *Journal* articles by Jon Sudbø et al. We asked each author to respond before September 30, 2006. We have now received responses from all authors except Asle Sudbø. Each of the responding authors, except Jon Sudbø, has indicated that the data that form the foundation for the articles have been called into question by the findings of the commission, and each of them has requested that the articles be retracted. Jon Sudbø alone does not agree with the commission's report. Given the weight of evidence offered in the commission's report and the requests of most of the authors of the articles, we retract both articles.

Retraction: Sudbø J et al. DNA Content as a Prognostic Marker in Patients with Oral Leukoplakia. N Engl J Med 2001;344:1270-8 and Sudbø J et al. The Influence of Resection and Aneuploidy on Mortality in Oral Leukoplakia. N Engl J Med 2004;350:1405-13

Gregory D. Curfman, M.D., Stephen Morrissey, Ph.D., and Jeffrey M. Drazen, M.D.

**Figure 3**

SUDBØS MOST CITED ARTICLE WAS RETRACTED IN NOVEMBER 2006[3]

[3] (Curfman , Morrissey , & Drazen 2006b)

In November 2006 the article was officially retracted on the request of all of the co-authors, only opposed by Jon Sudbø who did not agree with the findings of the Sudbø-Commission (Curfman et al., 2006b).

The committee focused its examination on the article because they considered it the most important article of Sudbø's career, and it was the basis for several subsequent articles, as well as a similar article published in the same journal in 2004. The focus of Sudbø's study was that by doing a ploidy analysis and comparing it with cancer development Sudbø hypothesized that one could use the ploidy in cells from white patches as an indicator of future oral cancer; a very serious form of cancer. The result of Sudbø's research had left him with an impressively good method to predict oral cancer for people with white patches. The extensive discrepancies found by the Commission lead them to conclude that: "*this sensational research result no longer have credibility*" (Ekbom et al., 2006, p. 81).

Sudbø's study was based upon patient material consisting of 150 patients. The findings of the committee did however revealed several fundamental problems with the *New England Journal of Medicine* 2001 and 2004, as well as a range of other articles:

- The same patient occurred multiple times. The mentioned 150 patients was at most 141, though it was difficult for the Commission to determine the actual number.
- A large number of patients should have been excluded from the study, since they already had existing or prior cases of oral cancer. At least 69 people should have been excluded.
- Ploidy analysis was not made for all patients. The Commission was only able to retrieve ploidy analysis from 69 of the observations, stemming from 65 different persons. They estimate that over 150 ploidy analysis had been performed, but the same patients appeared several times.
- The age distribution of the material did not correspond to each other from article to article. In the NEJM article from 2001 it claims that the material dates back to 1982 to 1995, and in other articles he claims that it dates back to 1976 to 1995.

The Commission viewed the failure to exclude patients from the study as the most serious one. When over half of the patient material should have been excluded from the study, the results of the study was no longer valid. Therefore all future research conducted by

Sudbø, based upon this material was also invalid. All publications based upon this research should therefore be retracted (Ekbom et al., 2006, pp. 58, 86-88).

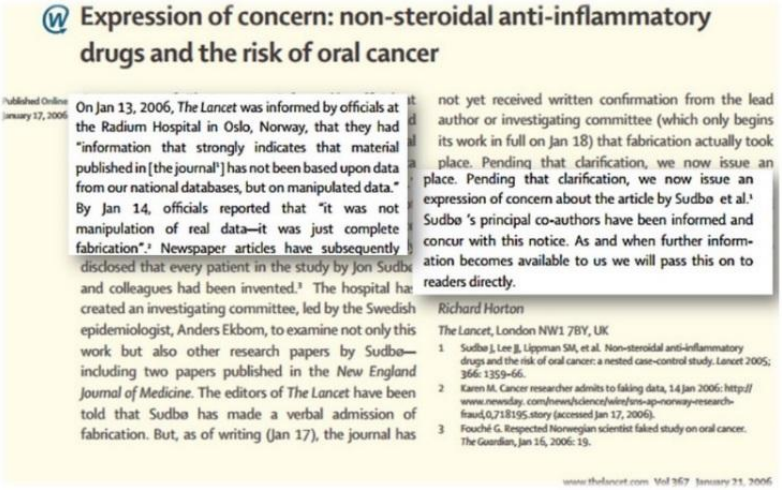
## 4.3 DISTRIBUTING THE BLAME

The Commission criticized several people for not catching Sudbø, ultimately blaming the institutions and their routines for letting Sudbø get away with his misconduct for so long. The media was quick to question the role of peer-reviewers and asked why they had not done their job.

Once people started to take a closer look at *the Lancet* article the cracks where easily spotted. Which led many to question the work of the peer-reviewers. When talking about the Sudbø-fraud in 2006, Paul Gerber use the Sudbø-case as an example of how scientific frauds sometimes can slip past peer reviewers. A reason he gives is the fact that is the inherent trust in the data, which a reviewer must display, since, like in the case of Sudbø, the data involves a specialized area of medicine and to replicate the data would be time-consuming and unrealistic (Gerber, 2006). In terms of *the Lancet* article, all fault for not discovering the fraud does not lie with the peer-reviewers. One of the peer-reviewers was the Danish professor Henrik Toft Sørensen, who worked as the Head of Department of Clinical Epidemiology at Aarhus University Hospital. He did in fact warn the editor of *the Lancet*, Richard Horton, against Sudbø's article, only to be ignored. The impression Sørensen had after reading the article was that something was off. Sørensen noticed that the article referred to data registries which were not available at the time of Sudbø's study. *"I had strong reservations against the quality of the article"* and his "*alarm bells were ringing*" Sørensen said to *forskning.no* when discussing the case back in 2011 (Hildebrandt, 2011). Ultimately Sørensen recommended that *the Lancet* should not publish the article. The editor, Horton, chose to ignore that recommendation. Something he easily glossed over when asked how *the Lancet* had not caught Sudbø's manipulations. He claimed that the peer-reviewers had unanimously approved the article for publishing. (Hildebrandt, 2011) Horton instead blamed the co-authors who he claimed they had had extensive contact with about the data and the analysis. Painting a picture of co-authors who were far more active than they actually had

been, claiming that it was "*incomprehensible how not one of the co-authors knew something*" (Hafstad, 2006).

On the day before the Sudbø-commission began their work, Horton published an expression of concern about *the Lancet* article. Before that he had marked the case of Sudbø as "*the biggest scam from a scientist the world had ever seen.*" Before concluding that *the Lancet* had followed all of the routines to ensure quality. (Hafstad, 2006). Many feel that this was his way of taking the spotlight away from himself, in order to distribute the blame elsewhere. However blaming the peer-reviewers, or the co-authors did not exonerate him completely, since he glossed over the fact that *the Lancet* had been in a hurry to publish the article. Michael Thun, an epidemiologist at the American Cancer Society was another person who had advice against the article, because he was not given enough time to check some of the discrepancies he had discovered with the article. Horton later denied any claims that he had been warned, accepting no responsibility (Tunstad, 2011).



Expression of concern: non-steroidal anti-inflammatory drugs and the risk of oral cancer

On Jan 13, 2006, *The Lancet* was informed by officials at the Radium Hospital in Oslo, Norway, that they had "information that strongly indicates that material published in [the journal[1]] has not been based upon data from our national databases, but on manipulated data." By Jan 14, officials reported that "it was not manipulation of real data—it was just complete fabrication".[2] Newspaper articles have subsequently disclosed that every patient in the study by Jon Sudbø and colleagues had been invented.[3] The hospital has created an investigating committee, led by the Swedish epidemiologist, Anders Ekbom, to examine not only this work but also other research papers by Sudbø—including two papers published in the *New England Journal of Medicine*. The editors of *The Lancet* have been told that Sudbø has made a verbal admission of fabrication. But, as of writing (Jan 17), the journal has not yet received written confirmation from the lead author or investigating committee (which only begins its work in full on Jan 18) that fabrication actually took place. Pending that clarification, we now issue an expression of concern about the article by Sudbø et al.[1] Sudbø 's principal co-authors have been informed and concur with this notice. As and when further information becomes available to us we will pass this on to readers directly.

*Richard Horton*

*The Lancet*, London NW1 7BY, UK

1 Sudbø J, Lee JJ, Lippman SM, et al. Non-steroidal anti-inflammatory drugs and the risk of oral cancer: a nested case-control study. *Lancet* 2005; 366: 1359–66.

2 Karen M. Cancer researcher admits to faking data, 14 Jan 2006: http://www.newsday.com/news/science/wire/sns-ap-norway-research-fraud,0,718195.story (accessed Jan 17, 2006).

3 Fouché G. Respected Norwegian scientist faked study on oral cancer. *The Guardian*, Jan 16, 2006: 19.

**Figure 4**

HORTON QUICKLY PUBLISHED AN EXPRESSION OF CONCERN IN *THE LANCET*[4]

Horton did however have a point, because the scientific fraud could have been dealt with sooner. Already in 2001 when the NEJM article (Sudbø et al., 2001) was first published a duplication of microscopy pictures was discovered in the article. Jon Sudbø explained that it was an honest error, and the mistake was not corrected or reported to the journal. If more action had been taken back then, the scandal that followed in 2006, might have been avoided (Oksvold, 2016). Sheldon Kotzin use this to point out that if an author claims that their instance of fraud is a one-time thing, a common error or an isolated event, that is not good

[4] (Horton, 2006)

enough and will require further investigation. "*You can't afford to give a scientist the benefit of the doubt, even someone of Sudbø's status*…"(Kotzin, 2007).

Erik Tunstad writes that the only person to be blamed for the misconduct is Jon Sudbø himself, condensing the misconduct to the fact that some people are *rotten apples* and even scientist sometimes cheat (Tunstad, 2011, pp. 10-11). The Sudbø-committee blamed the institutions and Sudbø's advisor, Albrecht Reith, for not having the proper routines and control of what Sudbø was doing. Horton blamed everyone, but himself, yet the Sudbø-committee concluded that Sudbø acted alone, and the co-authors had little to no knowledge of what was going on. The journals should have caught the misconduct, but the rushed treatment of *the Lancet* article proved that the peer-review is not always a watertight system, and sometimes articles that should not be published are. The purpose of retractions is to act as a corrective measure, but for that to have effect retractions must be handled correctly.

## 4.4 RESULT: HOW IS SUDBØ CITED?

The purpose of the citation count and analysis was to determine how two of Sudbø's retracted articles were cited after retraction. The two articles in question was his most cited one, and the one that lead to Sudbø being caught.

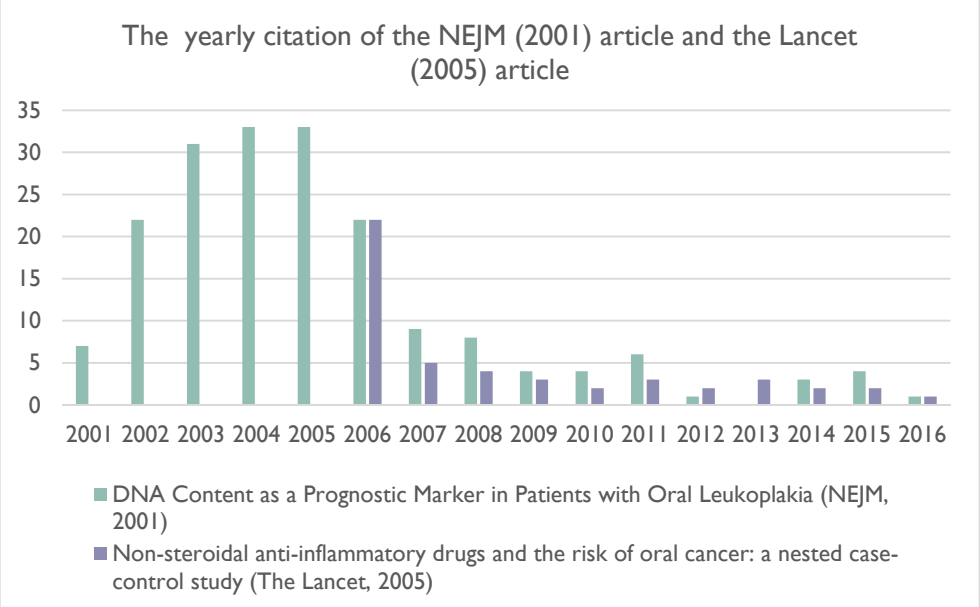



**Figure 5**

### 4.4.1 CITATION REPORT OF THE TWO ARTICLES

The *Citation report* from *Web of Science* revealed that Sudbø is still cited, with one citation for each of the articles in 2016, 10 years after the retraction. In fact the only year the NEJM (Sudbø et al., 2001) article did not receive any citations was in 2013. His most cited article from NEJM (Sudbø et al., 2001), had a predictable drop in citations after 2006 and the retraction. Yet the article has an average of 4 yearly citations after retraction, compared to the average of 24.7 citations before retraction. The article from *the Lancet (Sudbo et al., 2005)*, which exposed Sudbø shortly after being published, received most of its citations in 2006. This is to be expected, since the article had very little time to gain any traction before it was exposed, compared to the NEJM article. This goes hand in hand with the results of the citation count, which revealed that most of the citations connected to *the Lancet* article, were in terms of debate, and reporting about the fraud, using Sudbø as an example of scientific misconduct.

In total the two articles are cited 237 times in *Web of Science*, with the exclusion of one citation, because it linked to nothing, leaving the total count to 236 citations. Out of these citations 83 citations for the two articles combined, came after retraction.

### 4.4.2 WEB OF SCIENCE AND GOOGLE SCHOLAR CITATION COUNT

The NEJM article is cited 40 times after retraction, and *the Lancet* is cited 43 times, when you include the citations from 2006, which falls into the *grey area*.

In *Web of Science* the *NEJM* article were cited in the debates ($n$ = 4), does not mention retraction ($n$ =16), mentions retraction ($n$ = 5), grey are ($n$ = 6), and the no access category ($n$ = 5). *The Lancet* article was cited in debate ($n$ = 32), does not mention retraction ($n$ =3), grey area ($n$ = 4), and the no access category ($n$ = 1). In *Google Scholar* the citation count was higher with the total number of citations for both articles at 429 times. The *NEJM* article was cited 327 times, and *the Lancet* article was cited 102. After retraction the combined number of citations were 160, with the *NEJM* article cited 104 times and *the Lancet* article 56 times. The citations came from a number of sources, but the majority was from peer-reviewed journal articles. The *NEJM* article were cited in the debates ($n$ = 10), does not mention retraction ($n$ =31), mentions retraction ($n$ = 20), grey area ($n$ = 10) as well as places that I did not have

access to ($n$ = 33). *The Lancet* article was however only cited in debate ($n$ = 33), does not mention retraction ($n$ = 4), grey area ($n$ = 2), and the ones inaccessible to me ($n$ = 17), with no citations in the mentions retraction category ($n$ = 0).

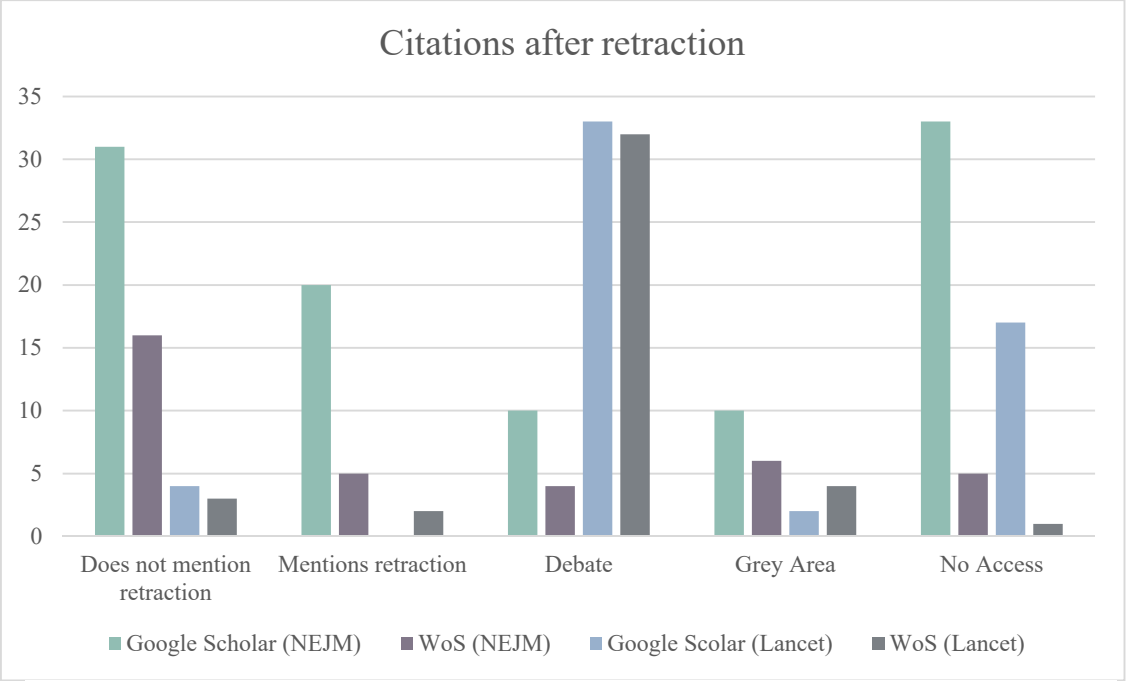



**Figure 6**

The results from the *Web of Science* search and the *Google Scholar* search are similar, but *Google Scholar* did of course yield a larger number of citations. The one that comes closest to each other is *the Lancet* article, who received almost the same amount of citations in *Google Scholar* and *Web of Science* in the debate category. Also notable is the amount of citations which I could not access in *Google Scholar* compared to *Web of Science*. This is to be expected since *Web of Science* indexes scientific journals, compared to *Google Scholar* who indexes much more broadly. Since *Thomson Reuters* only indexes journals that are members of ISI, the credibility of the citing sources is generally considered to be higher than from *Google Scholar*. Other than that the main difference between the *Web of Science* citation count and the *Google Scholar* one is in the percentage distribution of citations.

In *Web of Science* 44% of the citations of the NEJM article, did not mention retraction, compared to the 14% that did. An overwhelming amount of the citations of *the Lancet* article, 76%, were citations that talked about the Sudbø-fraud, commenting on his misconduct. The NEJM article is less used to exemplify misconduct, with only 11% using it for debate. This is to be expected since the publicity surrounding *the Lancet* article was much higher. This might also be the reason to why 17% of the NEJM citations falls into the grey area, where people had the time to include a comment about the Sudbø-incident but did not. Only 10% of *the Lancet* articles citations falls into the grey area.

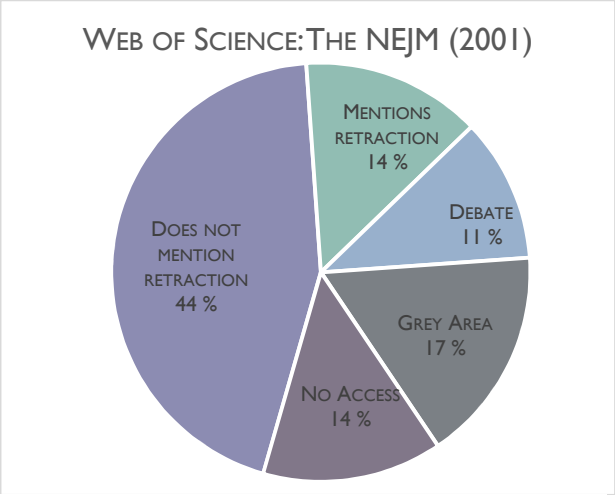



**Figure 7**

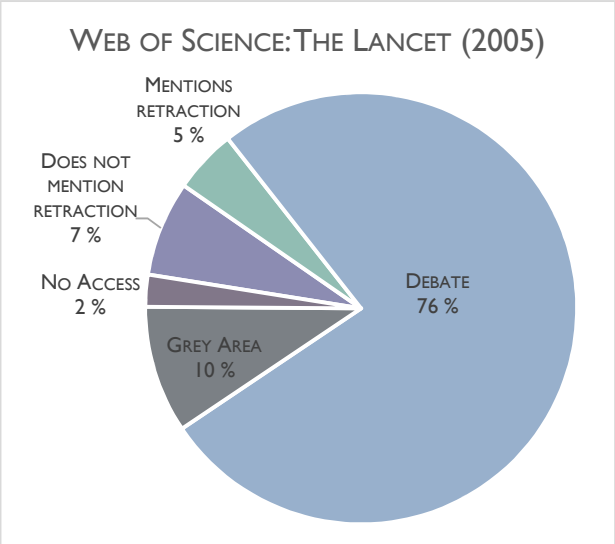



**Figure 8**

From *Google Scholar* 30% of the ones who cite the NEJM article, does not mention the retraction. 19% does mention retraction, and 9% use the article to talk about the fraud or debate Sudbø. Many of those debating Sudbø talked about the photo fraud, and how one can proceed to avoid photo manipulation in the future, since one of the problems with the NEJM article was the fact that there was photoduplication, with the same image appearing twice, but reported as different images. As with the search from *Web of Science*, there is a clear difference between the citing of the two articles. With *the Lancet* article 59% of the citations received after retraction is due to people talking about the fraud and misconduct. In total *the Lancet* article was cited 102 times, which means that *the Lancet* article is received more citations after retraction. Of the citations of *the Lancet* article only 4% of them were so-called "grey are" citations. This could be the result of the extreme spotlight which was on *the Lancet* article, with Horton issuing an expression of

concern almost immediately after the news broke. The expression of concern for the NEJM article was published in February 2006 (Curfman , Morrissey , & Drazen 2006a), with New England Journal of Medicine officially retracting the article in November 2006, on the request of the co-authors, which means that many might have been waiting for the conclusion of the Sudbø-committee, before taking action (Curfman et al., 2006b). This might explain why a larger percentage of the citations of the NEJM article falls into the "grey area", with 17% from *Web of Science*, and 10% from *Google Scholar*. In fact even after retraction more people still cite the NEJM article as a legitimate source, without mentioning the retraction. Which could be a result of people simply not being as aware of the retraction of the NEJM article as *the Lancet* article. There was quite a large number of articles which I did not have access to mainly due to language barriers. Therefore the result might have looked differently if I had known the content of these articles. The different languages citing Sudbø ranges from Italian, Dutch, French, Spanish, German and several Asian languages. I have no doubt that some of these sources are citing Sudbø to talk about his misconduct, but some might have been using him as a legitimate source.

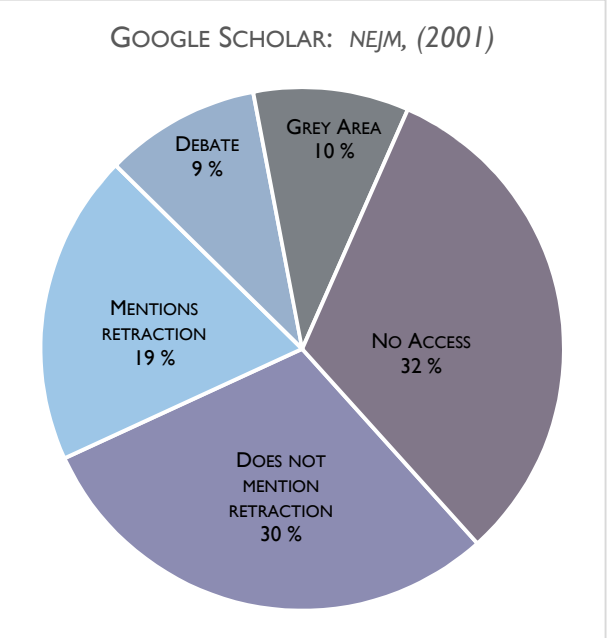



**Figure 9**

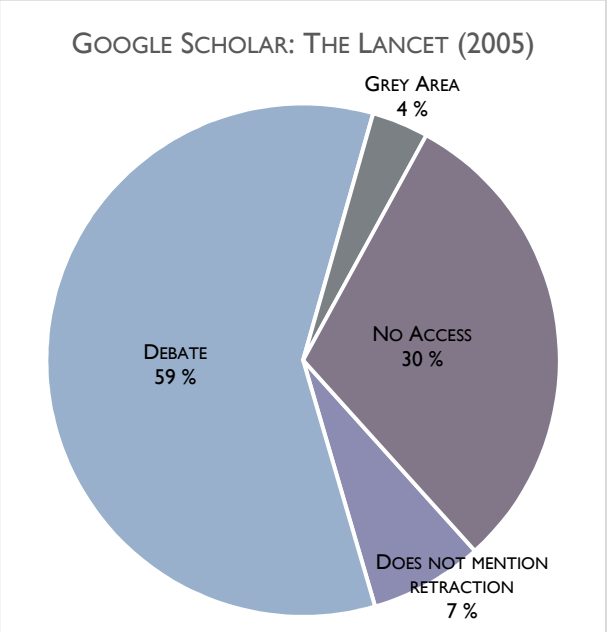



**Figure 10**

### 4.4.3 RETRACTIONS IN DIFFERENT DATABASES

If you look past the larger amount of press that *the Lancet* article received there might be other factors who determine the citation count, such as the handling of the retraction in different databases. The search through the different databases revealed that not everyone were handling the retractions according to the COPE-guidelines. When searching through *Web of Science*, *Google Scholar*, *PubMed*, *Scopus*, *ProQuest*, and *Medline*, all of the databases, except for *Google Scholar*, noted in the article search record that the two articles were retracted. This notice included any indication within the database search record that the article was retracted. However only *Web of Science* noted in the search record title that the articles were retracted. *Google Scholar* only did this for *the Lancet* article. Fortunately with the *NEJM* article all of the links to the PDF version showed that the article was retracted, but with *the Lancet* article PubMed, Scopus and Medline linked to PDF versions that did not show that the article was retracted. In Scopus you have the option to go to the publishers' version of the article, this takes you directly to the journals version, and both *The Lancet* and New England Journal of Medicine have published retractions to the articles. Since it is indicated in the search record for all, but *Google Scholar*, with the retractions and expressions of concern included in the search result, you would have to have pretty severe tunnel vision to make your way to the PDF.

| | **Noted in Article Search record** | | **Noted in Search Record Title** | | **Link shows retraction (PDF)** | |
|---|---|---|---|---|---|---|
| **Database** | *NEJM (2001)* | *The Lancet (2005)* | *NEJM (2001)* | *The Lancet (2005)* | *NEJM (2001)* | *The Lancet (2005)* |
| *Web of Science* | √ | √ | √ | √ | N/A | N/A |
| *Google Scholar* | X | √ | X | √ | √ | √ |
| *PubMed* | √ | √ | X | X | √ | N/A |
| *Scopus* | √ | √ | X | X | √ | X |
| *ProQuest* | √ | √ | X | X | √ | X |
| *Medline* | √ | √ | X | X | √ | X |

***X: No, √: Yes, N/A: not available.*** **Figure 11**

The redeeming factor for *Google Scholar* is the fact that the results link directly to the journals site, and they have posted the retraction. If you however only look at the article, without glancing at any of the other results you do have a small chance of making your way to the PDF version of the article, and in the case of *the Lancet* article this means that you might miss the retraction. If you have a hardcopy if the article and are only doing a preliminary check to look for retractions, you are less likely to make your way to the PDF-version, therefor you might miss the fact that the articles are retracted in Google Scholar. In other words for these two article you would be best served to use Web of Science to determine the retraction status, if you are not familiar with it from before.

# 5 DISCUSSION: WHY ARE RETRACTED ARTICLES CITED?

The result of the citation count from Sudbø's two articles, shows that sometimes it is necessary to cite retracted articles. Many of the citations for *the Lancet* article falls into the *debate* area. From the *Web of Science* citation count we can see that 76% of those who cite it use it to talk about the misconduct. If we were to apply the same reasons for citing as Brooks did, this would mean that a lot of the citations falls into the *negative credit* citing category. A category Brooks experienced that very few citer would put their citing into (Brooks, 1985). The crucial difference here might be the fact that the content of Sudbø's articles have been proven to be wrong, which was the whole reason for the retraction. So one of the most prominent services these article can provides us with today is the ability to learn from them, not from the content, but learning what not to do. The point of retractions is to correct the science. Articles that are retracted are generally labelled as having little value, and ideally they should fade into obscurity. By using it to talk about scientific misconduct you are giving the article a new value, which means that this type of *negative credit* citing is a positive thing.

The problem are the citations who cite after retraction without mentioning the retraction. The results of the Web of Science citation count showed that as much as 44% of the ones citing the *NEJM* article cite it without mentioning retraction. In other words they largely use it as *persuasion*. They cite Sudbø in order to support their own views and findings. This is the kind of citing that a retraction is meant to discourage.

Yet the citations for Sudbø did drop after citation. The *NEJM* article had an average of 24.7 citations before the retraction, and an average of about 4 yearly citations after retraction. *The Lancet* article was not cited until after the scandal broke, the difference in how the two articles are cited, both in the *Web of Science* count and the *Google Scholar* count, are therefore to be expected. The high publicity *the Lancet* article received may lend itself directly to ensuring that it is not cited in the same way as the *NEJM* article.

## 5.1 REASONS FOR CITING RETRACTED ARTICLES

When going through the literature some primary reasons for citing retracted articles stood out. One main issue is the importance of making the retractions known (van der Vet & Nijveen, 2016). This means to follow the COPE-guidelines for retraction, since significantly fewer retracted articles with the retraction notice freely accessible were cited, then articles where the retraction notice was not easily available (Balhara & Mishra, 2014). In the case of Sudbø's retracted articles this could be part of the problem. Both the *New England Journal of Medicine* and *the Lancet* have issued retractions according to the COPE-guidelines. The retractions are clearly defined as retractions and they are linked to the original article, who is marked as retracted. The retractions were published promptly on the website of the journals in order to minimize harmful effects. It is clearly stated who is retracting the articles, with the reason for retraction stated as well. However as we can see from my search through the databases, not all of the databases are in compliance with the guidelines. Only *Web of Science* ticks every box as they had noted of the retraction in the article search record, and search record title. The link they provide you with is to the journals version of the article, which is clearly marked as retracted. Only *Google Scholar* failed to note of the retraction in the article search record for the *NEJM* article. This is the article that is most cited without acknowledging retraction. As many as 44% have managed to miss the fact that the article is no longer part of the chronicles of science. An excuse for this can perhaps be found in the databases handling of the retraction. The *NEJM* article was published back in 2001. That is 15 years ago. The idea that some scientist have filed away a hardcopy of the article, in order to use it in the future is not so farfetched. Since the retraction warnings are added inconsistently in the databases, said scientist may have searched for the article in say *Google Scholar*, only to not be warned that the article is retracted, and is therefore using it as valid research which they can use in their *persuasive* citing (van der Vet & Nijveen, 2016), (Brooks, 1985). This incomplete, uninformative retraction notice can be a direct cause for the citations of retracted articles. One must also accept the possibility that if most of the people who cite do it for *persuasiveness* they might not be to inclined to look for evidence of retractions once they have found a source which describes something they agree with. A quote from George Bernard Shaw may illustrate my point:

*The moment we want to believe something, we suddenly see all the arguments for it, and become blind to the arguments against it*. (Shaw, 1928)

So if you have filed away a hardcopy of something, for future research, you might be content with that, and not bother to look for retractions. Or perhaps scientist are too trusting, believing in the honesty of others, and like Fang and Casadevall said: "m*ost scientist feel that research misconduct is uncommon*" (Ferric C. Fang & Casadevall, 2011).

One of Garfield's intended uses for the citation index was to avoid the uncritical citation of wrongful data, by letting the criticism of earlier papers be made available (Garfield, 2006a). Relevant papers across disciplines would be made available, where before it had been limited to one scientific area by only using subject or keywords (Yancey, 2005). This citing across fields may be a reason for the citing of retracted articles, since people who are not in the field, might not be aware of the research-scandals in a field. When Bornemann-Cimenti et al. researched the retractions of Reuben's, they found that the percentage of citations that cited from Reuben's field, varied between 36% to 60% in the years after retraction (Bornemann-Cimenti et al., 2015).

Bavelas talked about one motive for citing as the act of demonstrating your knowledge of the field (Bavelas, 1978). By citing articles that have been retracted what you are really doing is demonstrating your lack of knowledge of the field. She also points out that sometimes this can lead to the citing of articles that one has not even read (Bavelas, 1978). Simkin and Roychowdhury used stochastic methods to determine that as many as 70-90% of scientific citations were copied from other reference list (Simkin & Roychowdhury, 2005). If that is a common practice, than the citation of a retracted article is not an impossibility. If you combine this with Bavelas expectation that some people use sources that they have not read because it is expected of you, this could be a reason for the citation of retracted articles. However what Bavelas suggests are sources who have been established as so important to the field, be it from arbitrary rules, which means that this process requires some knowledge of the field (Bavelas, 1978). This knowledge should be extensive enough to weed out articles that are retracted, especially from something as publicised as the Sudbø-case was. Copying a reference is not the same as not reading it either. The problem of, in terms of retracted work, copying a reference only comes if that means that you do not take the time to read the original article, who will hopefully have been properly marked, making you aware of its retraction.

## 5.2 HOW CAN WE STOP IT?

Throughout the literature there have been a clear agreement that even though we are experiencing an increase in retractions, not nearly enough articles are being retracted. In the last few years articles retracted for research fraud have increased nearly tenfold (F. C. Fang et al., 2012). Some blame the increased pressure to publish for the increase of misconduct. The focus on the impact factor is seen as a negative thing by some, and is believed to be one of the underlying causes for the increase in misconduct, since so much rely upon it. Today the higher ranking journals advertise their impact factor in order to encourage people to publish in them. Jon Sudbø published in amongst others the New England Journals of Medicine and *the Lancet*, both high ranking journals. Yet his published articles were wrong.

> *"The Lancet has an Impact Factor of 45.217. The journal is currently ranked second out of 150 journals in the general medicine category"*
>
> (Lancet, 2016)

> *"The most recent (2014) impact factor for NEJM is 55.873, the highest among general medical journals."*
>
> (NEJM, 2016)

Bornemann-Cimenti et al. bring up the reviewing process, and the expectation put upon them to catch mistakes, such as the citation of a retracted article. With the NEJM article, one could expect it since people found errors with the article already back in 2001, so one could certainly question the peer-reviewing process. With the Lancet article, it was rushed, and the peer-reviewers were not heard, which means that not only the peer-reviews has a responsibility to look for errors, but also the editors. Reviewers and editors may be experienced researcher, yet they are not immune to making the same mistakes, or in this case, not catching the same mistakes that authors were unable to catch (Bornemann-Cimenti et al., 2015).

Fang and Casadevall suggest that there is a correlation between a high impact factor and a high retraction rate (Ferric C. Fang & Casadevall, 2011). Which serves as a reason to question the impact factor as a measure of quality. This focus on impact factor caused several Brazilian journals to do organized citation stacking in order to inflate their journals impact factors. They were eventually caught and blamed the focus on the impact factor for their

actions. (Noorden, 2013) The pressure to publish in higher ranking journals was an excuse Sudbø also used for his misconduct, claiming that it was his strive for academic greatness that caused him to do it.

The focus on bibliometrics and the use of citation counting have been debated over the years, some are eagerly for it, like Garfield, whilst others are against. Citation counting and the impact factor might not be the right tool to evaluate scientist, but in terms of the citing of retracted articles it is a good tool to better understand the phenomenon, mostly because it shines a spotlight on the fact that articles that are retracted are still cited. It provides a clearer understanding of the citation of retracted articles, which is important in order to devise ways to stop the citing. What has stood out to me as the best ways to do so is to increase the visibility of retractions. It is important to make people aware of them. Visibility might be one of the reasons to why higher ranking journals have a higher retraction rate, because they are put under a higher level of scrutiny, which might mean that misconduct is easier to spot (Ferric C. Fang & Casadevall, 2011). Retractions should be easily spotted to, which puts a heavy responsibility on the databases to make it clear if and why an article is retracted. As we have seen with the Sudbø-case, such a well-publicised case, some of the databases failed to make it abundantly clear that the articles were retracted. Another point is the development of automated processes which can check this for us. Such a computer program does not exist yet. Therefore the responsibility falls on us, to maintain constant vigilance. Blogs such as *Retraction Watch*, is an asset, because it makes people aware of scientific retractions across fields.

### 5.2.1 WHERE IS SUDBØ NOW?

Sudbø as a case exemplified well that retracted articles are cited. The level of misconduct that he showed, would make you think that he would not be. Yet his findings lives on through the wrongfull citations of his retracted articles.

And as for Sudbø himself? Today he is working as a dentist, banned from doing research. Yet he still manages to stir up controversy and strong feeling in others to this day. Just recently his image could be found in the pages of the Norwegian newspaper Aftenposten. They could report that he has been hired as leader of the clinic at Tannhelsetjenestens kompetansesenter Øst (Oral Care Competence Centre East). The University of Oslo found the

hiring of Sudbø to be so problematic that they are considering not sending their students there for training. The centre has emphasised to the Norwegian Institute of Public Health and to Nordic Institute of Dental Materials that Sudbø will in no way be involved in doing research, be it the collection of data or in application processes. According to the director of the Oral Care Centre, Hilde Vogt Toven, Jon Sudbø was hired, simply because he was the best applicant. (Dommerud, 2016)

# 6 LITTERATURE


Andersen, S. S. (2013). *Casestudier : forskningsstrategi, generalisering og forklaring* (2. utg. ed.). Bergen: Fagbokforl.

Balhara, Y. P. S., & Mishra, A. (2014). Compliance of retraction notices for retracted articles on mental disorders with COPE guidelines on retraction. *Current Science, 107*(5), 757-760.

Bavelas, J. B. (1978). The social psychology of citations. *Canadian Psychological Review/Psychologie canadienne, 19*(2), 158-163. doi:10.1037/h0081472

Bellis, N. d. (2009). *Bibliometrics and citation analysis : from the science citation index to cybermetrics*. Lanham, Maryland: The Scarecrow Press.

Bornemann-Cimenti, H., Szilagyi, I. S., & Sandner-Kiesling, A. (2015). Perpetuation of Retracted Publications Using the Example of the Scott S. Reuben Case: Incidences, Reasons and Possible Improvements. *Science and Engineering Ethics*, 1-10. doi:10.1007/s11948-015-9680-y

Brooks, T. A. (1985). Private acts and public objects: An investigation of citer motivations. *Journal of the American Society for Information Science, 36*(4), 223-229. doi:10.1002/asi.4630360402

Bruhns, S. (2010). *Citationsindexering*. Romdrup: Bibliografica.

Budd, J. M., Sievert, M., Schultz, T. R., & Scoville, C. (1999). Effects of article retraction on citation and practice in medicine. *Bulletin of the Medical Library Association, 87*(4), 437-443.

Cokol, M., Iossifov, I., Rodriguez-Esteban, R., & Rzhetsky, A. (2007). How many scientific papers should be retracted? *Embo Reports, 8*(5), 422-423. doi:10.1038/sj.embor.7400970

COPE. (2009). COPE Committee on Publication Ethics: Retraction Guidelines. Retrieved from http://publicationethics.org/files/retraction%20guidelines_0.pdf

COPE. (2016). COPE Committee on Publication Ethics: About COPE. Retrieved from http://publicationethics.org/about

Curfman , G. D., Morrissey , S., & Drazen , J. M. (2006a). Expression of Concern: Sudbø J et al. DNA Content as a Prognostic Marker in Patients with Oral Leukoplakia. N Engl J Med 2001;344:1270-8 and Sudbø J et al. The Influence of Resection and Aneuploidy on Mortality in Oral Leukoplakia. N Engl J Med 2004;350:1405-13. *New England Journal of Medicine, 354*(6), 638-638. doi:doi:10.1056/NEJMe068020

Curfman , G. D., Morrissey , S., & Drazen , J. M. (2006b). Retraction: Sudbø J et al. DNA Content as a Prognostic Marker in Patients with Oral Leukoplakia. N Engl J Med 2001;344:1270-8 and Sudbø J et al. The Influence of Resection and Aneuploidy on Mortality in Oral Leukoplakia. N Engl J Med 2004;350:1405-13. *New England Journal of Medicine, 355*(18), 1927-1927. doi:doi:10.1056/NEJMe068247

Dommerud, T. (2016). Tatt for forskningsjuks - nå er Jon Sudbø ansatt som klinikkleder. *Aftenposten*. Retrieved from http://www.aftenposten.no/nyheter/iriks/Tatt-for-forskningsjuks---na-er-Jon-Sudbo-ansatt-som-klinikkleder-8462842.html

Ekbom, A., Helgesen, G. E. M., Tverdal, A., Lunde, T., & Vollset, S. E. (2006). *Report from the Investigation Commission appointed by Rikshospitalet – Radiumhospitalet MC and the University of Oslo January 18, 2006*. Retrieved from http://www.ous-research.no/general/docs/ekbom/Report_Investigation_Commission.pdf

Fang, F. C., & Casadevall, A. (2011). Retracted Science and the Retraction Index. *Infection and Immunity, 79*(10), 3855-3859. doi:10.1128/IAI.05661-11

Fang, F. C., Steen, R. G., & Casadevall, A. (2012). Misconduct accounts for the majority of retracted scientific publications. *Proc Natl Acad Sci U S A, 109*(42), 17028-17033. doi:10.1073/pnas.1212247109

Fukuhara, A., Matsuda, M., Nishizawa, M., Segawa, K., Tanaka, M., Kishimoto, K., . . . Shimomura, I. (2005). Visfatin: A Protein Secreted by Visceral Fat That Mimics the Effects of Insulin. *Science, 307*(5708), 426-430. doi:10.1126/science.1097243

Fulton, A. S., Coates, A. M., Williams, M. T., Howe, P. R. C., & Hill, A. M. (2015). Persistent Citation of the Only Published Randomised Controlled Trial of Omega-3 Supplementation in Chronic Obstructive Pulmonary Disease Six Years after Its Retraction. *Publications, 3*(1), 17-26. doi:10.3390/publications3010017

Garfield, E. (2006a). Citation indexes for science. A new dimension in documentation through association of ideas. *International Journal of Epidemiology, 35*(5), 1123-1127. doi:10.1093/ije/dyl189

Garfield, E. (2006b). THe history and meaning of the journal impact factor. *JAMA, 295*(1), 90-93. doi:10.1001/jama.295.1.90

Gerber, P. (2006). What can we learn from the Hwang and Sudbo affairs? *Medical Journal of Australia, 184*(12), 632-635.

Grieneisen, M. L., & Zhang, M. (2012). A comprehensive survey of retracted articles from the scholarly literature. *Plos One, 7*. doi:10.1371/journal.pone.0044118

Hafstad, A. (2006, 17.01). - Største svindel verden har sett. *Aftenposten*. Retrieved from http://www.aftenposten.no/nyheter/iriks/--Storste-svindel-verden-har-sett-6382117.html

Hildebrandt, S. (2011). Ignorerte advarsel om Sudbøs juks. Retrieved from http://forskning.no/forskningssvindel/2011/03/ignorerte-advarsel-om-sudbos-juks

Horton, R. (2006). Expression of concern: non-steroidal anti-inflammatory drugs and the risk of oral cancer. *Lancet, 367*(9506), 196-196. doi:10.1016/s0140-6736(06)68014-8

Jakobsen, H. S. (17.01.2014). Tellekantens Tyranni. *Morgenbladet,* pp. 8-11. Retrieved from http://morgenbladet.no/samfunn/2014/tellekantenes_tyranni

Kotzin, S. (2007). From Darsee to Sudbö: NLM's role in the retraction process. *From Darsee to Sudbö: NLM's role in the retraction process, 4*(1).

Lancet. (2016). The Lancet: about. Retrieved from http://www.thelancet.com/lancet/about

Liu, Y., & Rousseau, R. (2013). Interestingness and the essence of citation. *Journal of Documentation, 69*(4), 580-589. doi:doi:10.1108/JD-07-2012-0082

Merton, R. K. (1968). The Matthew Effect in Science. *Science, 159*(3810), 56-63.

NEJM. (2016). New England Journal of Medicine: Media Center Fact Sheet. Retrieved from http://www.nejm.org/page/media-center/fact-sheet

Noorden, R. V. (2013). Brazilian citation scheme outed. *Nature, 500*(7464). doi:10.1038/500510a

NTB. (18.01.2006). - Sudbø fikk pasienter over på farligere medisin. Retrieved from http://www.vg.no/forbruker/helse/helse-og-medisin/sudboe-fikk-pasienter-over-paa-farligere-medisin/a/300648/

Oksvold, M. P. (2016). Incidence of Data Duplications in a Randomly Selected Pool of Life Science Publications. *Science and Engineering Ethics, 22*(2), 487-496. doi:10.1007/s11948-015-9668-7

Oransky, I. (2015). Top 10 most highly cited retracted papers. Retrieved from http://retractionwatch.com/the-retraction-watch-leaderboard/top-10-most-highly-cited-retracted-papers/

Shaw, B. (1928). *The Intelligent Woman's Guide to Socialism and Capitalism*: Transaction Books.

Simkin, V. M., & Roychowdhury, P. V. (2005). Stochastic modeling of citation slips. *Scientometrics, 62*(3), 367-384. doi:10.1007/s11192-005-0028-2
Steen, R. G. (2011). Retractions in the scientific literature: is the incidence of research fraud increasing? *Journal of Medical Ethics, 37*(4), 249-253. doi:10.1136/jme.2010.040923
Sudbo, J., Lee, J. J., Lippman, S. M., Mork, J., Sagen, S., Flatner, N., . . . Dannenberg, A. J. (2005). Non-steroidal anti-inflammatory drugs and the risk of oral cancer: a nested case-control study (Retracted Article. See vol 367, pg 382, 2006). *Lancet, 366*(9494), 1359-1366. doi:10.1016/s0140-6736(05)67488-0
Sudbø, J., Kildal, W., Risberg, B., Koppang, H. S., Danielsen, H. E., & Reith, A. (2001). DNA Content as a Prognostic Marker in Patients with Oral Leukoplakia. *The New England Journal of Medicine, 344*(17), 1270-1278. doi:10.1056/NEJM200104263441702
Tunstad, E. (2011). *Juks : hvordan forskere svindler - og hvorfor det ikke er så farlig*. Oslo: Humanist forl.
van der Vet, P. E., & Nijveen, H. (2016). Propagation of errors in citation networks: a study involving the entire citation network of a widely cited paper published in, and later retracted from, the journal Nature. *Research Integrity and Peer Review, 1*(1), 1-10. doi:10.1186/s41073-016-0008-5
Yancey, R. (2005). Fifty years of citation indexing and analysis *Thomson Scientific*.

# 7 IMAGES AND FIGURES:



The cover page image is a collage screenshots from the different articles that are cited throughout the paper. It also include a picture from VG and the word *retracted*, which I have referenced below.

- The word Retracted: http://img.medscape.com/thumbnail_library/is_151104_stamp_retracted_800x600.jpg
- "Jeg har betalt prisen" (VG Helg, 09.05.2009)