Distribution Agreement

Signature:

March 30, 2021

____________________________ ______________

Kerryn Ashleigh Roome Date

Publication Bias:

Its Role in Medical Literature & the Ethical Implications

By

Kerryn Ashleigh Roome
Master of Arts

Bioethics

Kathy Kinlaw, MDiv
Advisor

Rebecca Pentz, PhD
Committee Member

Michael Goodman, MD, MPH
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

Publication Bias:

Its Role in Medical Literature & the Ethical Implications

By

Kerryn Ashleigh Roome
B.S., University of Georgia, 2016

Advisor: Kathy Kinlaw, MDiv

An abstract of
A thesis submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Master of Arts
in Bioethics
2021

Abstract

Publication Bias:

Its Role in Medical Literature & the Ethical Implications

By Kerryn Ashleigh Roome

Publication bias, also called positive-outcome bias and dissemination bias, has been a long standing and prominent problem in the scientific community. Publication occurs for several reasons, but this thesis aims to analyze the ethical implications of such bias in medical literature through the lens of several different moral theories – feminist ethics, the ethics of care, utilitarianism, and Kantian ethics. This thesis also explores the ethical obligations all stakeholders have in publishing in an ethically sound manner as well as identifying the powers at play in a competitive and high stakes environment. Additionally, since this specifically has not been quantitatively measured or assessed, this thesis quantifies the status of publication bias in today's medical society as well as determines the changes that have occurred over a 30-year time period. Every original, hypothesis-driven research article in the top 3 and bottom 3 medical journals, ranked by impact factor, are coded and assessed in the years 1998, 2008, and 2018. This sampling technique shows trends in publication bias and provides insight into the status of publication bias. It was found that non-significant studies were more likely to be published if they were experimental studies rather than observational studies. Additionally, a significant change over time was found within high impact journals and observational studies.

Publication Bias:

Its Role in Medical Literature & the Ethical Implications

By

Kerryn Ashleigh Roome
B.S., University of Georgia, 2016

Advisor: Kathy Kinlaw, MDiv

A thesis submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Master of Arts
in Bioethics
2021

Acknowledgements

First and foremost, I want to thank David Rink for his assistance in the statistical department of this thesis. His volunteering to act as the 2nd coder, as well as complete the statistical analysis will be forever treasured.

Additionally, I would like to thank Emory’s Center for Ethics for welcoming me into the family and teaching me how to be a better advocate for my future patients. I will carry this training and knowledge with me forever and utilize it whenever possible.

List of Figures and Tables

Table of Contents

I: AN INTRODUCTION

What is publication bias and why is it a problem?

> “Appearances to the mind are of four kinds.
>
> Things either are what they appear to be;
>
> Or they neither are, nor appear to be;
>
> Or they are not, and do not appear to be;
>
> Or they are not, and yet appear to be.
>
> Rightly to aim in all these cases
>
> Is the wise man’s task.”
>
> -Epictetus, 2nd century AD [1]

For decades and even centuries, the scientific community has been plagued by the problem of publication bias, also known as positive-outcome bias and dissemination bias. The first mention of publication bias can be traced back to 1661; the chemist, Robert Boyle, stated that “many excellent notions or experiments are, by sober and modest men, suppressed” [2]. Over the years, it has garnered more and more attention as a significant problem, with the number of references to publication bias increasing from 37 between 1985 and 1989 to 2071 between 2010-2012 [3]. Moreover, as increased attention has been given to publication bias, there has been an active effort to implement safeguards to keep it under control. For example, in 2016, a statistical model (Bayesian model) was shown to provide assistance in truly understanding the impact a study is making thus lending itself well to alleviating publication bias [4].

Publication bias has been defined as a phenomenon in research in which studies with statistically significant results are more likely to be submitted to journals for review and are more likely to subsequently be published than studies with null or non-significant results [5, 6]. Additionally, publication bias occurs when studies that do not align with the original hypothesis are not submitted for review or are not considered for publication by journal review boards. As stated above, studies with positive outcomes are more likely to be published in journals and thus reach the scientific community; conversely negative studies which are not published do not get disseminated, potentially hampering research [7-11]. It has also been found that studies with statistically significant findings are published more quickly than those negative studies [12]. This devalues negative outcomes and may force the hand of researchers to only choose research topics that they are confident will lead to positive outcomes, and thus a greater number of publications.

Publication bias can occur from two main sources – the reviewers/editors of journals and the researchers themselves. Due to outside pressures such as revenue generated from subscriptions, reputation, and academic requirements, researchers may not even submit studies for review if the results are negative as they assume they will not be published. This is called the file-drawer effect and has been noted frequently. “Publication bias distorts the accumulated data in the literature, causes the over estimation of potential benefits of intervention and mantles the risks and adverse effects, and creates a barrier to assessing the clinical utility of drugs as well as evaluating the long-term safety of medical interventions” [13]. Since published literature is the source of health-policy data which forms the basis for the majority of clinical guidelines, the possibility of not having an accurate representation of research that has been done has the potential to be quite harmful, impacting the medical community’s ability to accurately synthesize and draw appropriate conclusions from the evidence [2, 3]. The impact of publication bias can vary widely, from the omission of one single study that would influence the overall understanding of a larger field to completely compromising a therapeutic area [14-16]. It cannot be reiterated enough the detriment that this practice causes – “distort[ing] the literature, drain[ing] scarce resources by undertaking research in futile quests, and lead[ing] to misguided research and teaching practices” [1].

Due to this being a well-known problem, clinical trials are now required to be registered so that reviewers have access to trials that were not published or not submitted for review [3]. While this is an important step forward, there is evidence that it is undermined using loopholes, lack of compliance, and poor enforcement by the agencies that oversee it [14-16]. It does not alter the information that ends up in journals that clinicians and scientists read and base decisions on. If the results of published studies are different from the results of unpublished studies, those studies that are published may very well not be an accurate representation of the collection of studies conducted [3]. Previous studies have estimated that approximately 50% of completed studies go unpublished [3]. This tendency in research is unethical and dangerous to many entities for several reasons, which will be explored in depth in this thesis. Since researchers have certain obligations to their field and to the public’s trust, how are these furthered or undermined by publication bias? What other stakeholders are involved in perpetuating publication bias, and what are their moral obligations?

There have been several reasons recorded for why publication bias occurs. Just like any other kind of bias, publication bias can occur “intentionally or unintentionally, consciously or unconsciously” [3]. Most commonly, it has been found that the reason for a completed study to go unpublished is due to the investigators foregoing writing a manuscript and failing to submit to a journal due to negative or nonsignificant results [3]. The highly competitive environment in which research lives only fuels this problem. For many, publishing is necessary for career promotion and funding – especially publishing new and exciting results. Due to this pressure to, it may force the hand of some [1]. Hooper also describes other reasons why completed results are unpublished such as journal rejection and a perception of unimportant results. High impact journals are notoriously difficult to publish in, and researchers may know what types of studies would be preferred. This may guide research projects as well as submissions for publication. Additionally, editors can play a role as they participate in a competitive environment as well and “it is the competition for citation index and the financial survival of journals that makes it more attractive to publish positive findings” [1].

While the academic community has become increasingly aware of the prevalence and dangers of publication bias, it appears that over the years there has been a continual and significant increase in publication bias [1]. The first study that provided evidence of publication bias was in 1959, by Sterling, where it was found that more than 95% of the articles in 4 prominent psychology journals reported statistically significant, or positive, results [2]. This study was repeated 29 years later, in 1988, and no changes were found [2]. A similar finding was confirmed in medical journals, with 85% of the articles reporting positive findings. Additionally, between 1990 and 2007, there was a statistically significant increase of 22% in the frequency of papers finding statistically significant positive results [1]. Could this simply be due to science improving and researchers becoming better at creating and testing hypotheses? Studies show this is highly unlikely and that studies resulting in only significant results is high unlikely as well [17]. Joober states that the poor replicability of study results makes it so.

In addition to an increase in publication bias over the years, it has also been found that there is a significant correlation between a journal's impact factor[1] and the effect size[2], with journals with higher impact factors overestimating effect sizes [1]. Are there additional variables that impact whether or not a study is more or less likely to be published to the extent that statistical significance does? Prior research has shown that other variables such as clinical versus observational trials, sample size, source of funding, and multicenter versus single center do not impact publication bias consistently [18]. The extent of publication bias is a threat to the validity and strength of ethically sound research, interpretation, and clinical practice.

---

[1] Impact factor is a common measure of perceived importance of a scientific journal. It is a numerical number given to journals based on how often they are cited by others compared to the total number of citable articles in said journal. Impact factor will be further described in Chapter 7.

[2] Effect size is a statistical measurement pertaining to the strength of a relationship between two variables. This numerical value allows one to draw conclusions about how much one variable impacts another, and vice versa.

As we move forward in the discussion of publication bias, it is important to note that the purpose of this thesis is not to argue that all studies are worthy of publication. The purpose of the arguments and points that will be made are to understand perhaps some of the reasons why publication bias may exist, how we can combat it, and what we should be aware of when thinking about the studies that are published in medical literature. There are many reasons why studies are rightly rejected for publication, and this thesis is not attempting to argue that all studies should be published. Rather, that we should perhaps be more thoughtful and aware of some of the faults that may exist in the current process of publication in medicine.

Structure

The structure of this thesis is multi-fold. We begin with a deep dive into the ethical obligations that parties involved in publishing medical literature have as well as the power struggle that may be present. Next, because there is not one moral theory that can fully explain or support any problem in society, we will explore several different arguments for why publication bias is unethical and posit several different explanations for its existence. The theories utilized include feminist and virtue ethics, utilitarianism, and Kantian ethics. Next, counter-arguments will be outlined as there may be several good reasons for why many significant studies are published over non-significant studies. After this exploration, the methods and results of a study to indirectly quantify publication bias will be addressed. Finally, the findings will be discussed and hypotheses for why certain findings were present will be posed. To conclude this thesis, we will explore steps to respond to publication bias and reflect on the importance of study findings, as well as how they can be utilized in the future to guide more ethically sound research and publication.

II: ETHICAL OBLIGATIONS AND THE POWERS AT PLAY

One explanation of why publication bias is so pervasive is power. How one defines power differs greatly, however the use of it can influence whether outcomes that are either right or wrong. Power has the ability to corrupt a moral identity or to enable a moral identity to arise [19]. There is the power one has over another, and then there is the power one has to accomplish or achieve a certain ideal. The former definition is one that Max Weber has said to be “the probability that one actor within a social relationship will be in a position to carry out his own will despite resistance” [20]. The latter, from Thomas Hobbes, states that power is a person’s “present means…to obtain some future apparent Good” [21]. I would argue these 2 can explain one another and walk hand in hand. As is the case with most ways to exist, there is not just one form of power. Power comes in many different forms. In the late 1950’s, social psychologists John French and Bertram Raven outlined 5 main types of power – reward power, coercive power, legitimate power, referent power, and expert power [22]. Here we will define these different versions of power and explore how each one could impact the occurrence of publication bias. The essence of power, no matter which definition you use, can explain the reason publication bias exists – or rather, the forces behind the reason publication bias exists.

French and Raven’s ultimate goal in delineating these five different types of power was to answer questions relating to the reasons behind the actions of one who holds and exerts power, compared to the reasons behind the reactions of those who endure the actions of the one who exerts the power. This would help further the understanding of different relationships in society, and in this case, can be useful to further understand how and why publication bias may occur. In order to fully grasp the meaning of these different forms of power, it should be noted that they were all defined from the perspective of the person whom power is exerted over [22]. It is important to remember that power is not always used to be “evil,” and that each type of power does not necessarily mean it is negative in nature. For arguments sake, however, these definitions will explore how they could be at fault, or rather, how each type of power could potentially be used to further publication bias. It is important moving forward to recognize how the different types of power impact the dynamics of publishing research, causing us to be aware of times when our actions are pivoting towards a negative use of power. If used correctly, several of these types of power could indeed be protective against publication bias.

(1) Reward power is the first of five, and speaks to the feeling that the person/entity in charge or with power has control over the other’s ability to receive a reward. In publication, an example could be found in the relationship between the larger organization that runs a journal (mainly editors and publishers), and the consumers of the research (i.e. readers and subscribers). In some ways, the consumers exert the power over the journal, the editors, and publishers. Without consumers, the journal would not survive. Thus, in essence, the content that is “sold” drives readership, subscriptions, and further citations (now with implications for impact factor). Ultimately, it could be that the readers have power over the editors and publishers, and thus, the journal’s content. Additionally, publishers can pressure editors to select certain papers in order to improve the journal’s notoriety, however many guidelines for editors clearly state that they are

to resist these pressures [23]. Whether or not that routinely occurs is unknown. This in turn has the ability to produce reward for those involved in the journal (increased pay, steady job, recognition, etc). To maximize this, the content will be carefully selected and curated.

(2) The second type of power is coercive power. It defines a power in which there is "an overt threat of harm intentionally presented by one person to another in order to obtain compliance" [24]. This could force the hand of someone, if they feel they would be better off pursuing the outcome the one in power would prefer. While a bit subtle and less obvious, this could potentially be seen between consumer (readers and subscribers of a journal) and the journal (editors, publishers) or between editors and publishers, however it seems to be a much less likely form of power that would be utilized to further publication bias. Coercive power is more likely to be found earlier on in research, most notably in situations where researchers exert their power over research participants – whether that be through payment or other incentives.

(3) The third type of power is legitimate power. This comes from the perception that the one with power is rightly owed that power and is appropriately in charge of the other. This type of power is oftentimes reasonable and well assigned, but blindly going along with an idea or set of actions simply because the other seems to be more knowledgeable can lead to problems. We must always remember to be curious and question certain things when they do not seem appropriate. An example of this in the process of publication bias is the relationship between reviewer and researcher. Researchers submit studies to be reviewed, oftentimes,

with the expectation that they will be rejected. The understood dichotomy is that the journal review boards have more power than the individual researcher and are rightly in charge of determining the fate of the submitted study. If a researcher feels the study they conducted is too weak, or does not provide impressive enough results, the study may be disregarded (file-drawer effect). This feeling of being powerless and premature defeat allows for publication bias to persist. Many may argue that legitimate power should not lead to publication bias, however legitimate power inherently stems from societal values and influence. If our values are not grounded in moral actions, such as just publication processes, then the "legitimate" power is flawed and encouraging the wrong behaviors. The impact of these will be further unpacked as we move forward in different ethical theories.

(4) The fourth type of power is referent power, which speaks to a less obvious form of power over someone. In referent power, the entity whom power is usually exerted over identifies with and relates to the other party involved. Due to this perceived likeness, they are more likely to go along with the other's plans or aspirations in order to maintain the relationship. In this kind of power, the influence one has over another is often missed even though the power exerted over them is strong. This could easily be imagined in committee meetings when discussing studies and determining their fit for the journal or perceived importance, and therefore likelihood to be published. This could occur in the research team when an esteemed member of the team believes that the study is

not worthy of publication, if other team members who disagree do not advocate for the study, it may very well not get published or even submitted for review.

(5) The fifth type of power is expert power, which is similar to referent power but instead of relating to those with the power, one feels as though the other party has more knowledge than they do and thus, doing what they desire is best. This could play out in publication processes and decisions in many ways – for example, if someone performs a study, but another party states the results are unimpressive, one may defer to the other's knowledge base and assume they are correct. This may lead them to not publish the results, or for it to be rejected for publication if the more knowledgeable party is that of the journal reviewers. Depending on the motives behind such characterizations, power can be utilized for selfish or just reasons.

While unrelated to publication bias, a parallel I feel sheds light on power within the system that can promote publication bias relates to physicians and their patients. Power is often a topic that is brought up when discussing patient-physician relationships. The power should be shared, physicians should recognize the power of the white coat and use their great influence to better the lives of others through that platform. I argue that this should be the same conversation had in research – a field, no different to any other, in which power dynamics are present. There are many stakeholders in research – research subjects, investigators, writers, reviewers, editors, publishers, readers, and we cannot forget funders. Within the category of readers, I would include those impacted downstream by the research – patients, healthcare providers, caregivers, professional healthcare associations. The external influences on each stakeholder carries immense weight.

The pressure to publish meaningful and interesting work is great, the community in which medical research exists is competitive, and the stakes are increasingly high. For this reason, it is easy to see why the less impressive studies are not submitted or are rejected for publication. While it has been extensively shown that publication bias exists, we must not forget the ethical obligations that each entity involved in the process of research holds. To ignore these obligations is to undermine the integrity of research and introduce unnecessary doubt and suspicion.

As alluded to previously, power can and has caused well-intentioned persons to pursue rather heinous outcomes. In fact, for decades, philosophers have argued power is closely tied with evil [25]. Power can encourage some to only serve themselves and ignore the common good. The goal of a philosopher is not to label an action or a person as “evil,” but rather to work to understand the meaning of evil and how that state comes to be [26]. How then, do we tie evil into the problem of publication bias?

Evil has been explored in many different ways – at both an individual and population level, as natural and moral, and as a form of suffering [26, 27]. All of these lend themselves well to the argument made when it comes to publication bias. In this paper the focus is on the effect of publication bias at the population level. Publication bias causes harm on a larger scale and this is where it relates to suffering. Claudia Card, a philosopher who based many of her theories on Kant, Schopenhaur, and Nietzsche, said that “evils are reasonably foreseeable intolerable harms produced (maintained, supported, tolerated, and so on) by culpable wrongdoing. So understood, evils have two irreducibly distinct components: a harm component and an agency component” [26]. When one could pursue different actions that could lead to less suffering and harm, but chooses not to for whatever reason, we could place those actions on the scale of evil. For when these choices are made, suffering is not equal, and falls primarily on a certain group. When one chooses to publish based on reputation or the “likeability” of a study outcome (i.e. the fit of the study in the journal’s theme, how much the study affirms current views, etc.), rather than the quality of research, harm and suffering may result.

To further this exploration of evil, natural and moral evil should be explained as it could provide a reference point for many arguments down the line [28]. The distinguishing features between these two kinds of evil, again, speak to the choice one has and the power that goes along with that. The existence and actions of moral agents is important in this as well. For the case of natural evils, it can be said that they are ills that occur by no direct doing of a moral agent [28]. Think of natural disasters and disease, for example. One could argue that moral agents do indeed play a role in these occurring – humans have long neglected the climate and health of the Earth and thus are suffering from a slew of natural disasters that may have been avoided. Additionally, someone with diabetes may have not exercised enough or eaten healthy foods, and thus could be responsible for their disease. To these arguments, I would agree, however for the sake of distinguishing these two distinct categories of evil, it is best to continue along the simple argument as it touches on the main differences – intentions and negligence. In contrast to natural evils, moral evils rely heavily on the actions of moral agents. On the part of a moral agent, if intentional actions are taken to further one's own agenda at the cost of others, it could be argued that these actions are evil in nature. Such intentional action may be taken by those involved in supporting and enabling publication bias.

In the setting of pervasive publication bias, it truly seems this is the case. One hypothesis could be that in an effort to only improve one's status in the medical community and receive additional funding, researchers do not submit their studies that "failed" or did not provide any significant results [29]. As evidence of this, in a groundbreaking paper published by Hwang, it was shown that amongst trials that were carried out over a 10-year time span, 54% of drugs/medications in late stages of clinical trials failed to be effective or safe, but only 40% of trial results were published for those failed drugs und investigation [29]. Hwang rightly brings to our attention the harm this failure to publish undesirable results can cause as it wastes funding and resources and halts the progression of medical advances [29]. In stark contrast, the rate of publication for trials that resulted in positive findings (namely, effective and safe drugs) has been found to be twice as high [30]. In an effort to better the reputation and impact of a particular journal, reviewers reject those studies that are not impressive or do not have positive findings. But to what detriment? This practice does not aid the medical community in understanding the literature fully. A piece of the puzzle is left out when there is bias. And when several pieces are missing, it is difficult to draw any conclusions. In the age of evidence-based medicine, conclusions and guidelines are based on the data presented in the literature and therefore if the data presented is incomplete, the validity of those conclusions and guidelines become highly questionable [2]. The published literature then informs further research endeavors and creates areas of interest that will receive more funding than others. This, in turn, creates additional pressure to conduct and publish studies that will only result in positive findings. The cycle continues, and becomes a vicious one.

Having discussed some aspects of power, we need to understand how, given a person’s moral identity, this may contribute to issues related to publication bias. Previous studies have investigated why some people are so quick to dismiss the common good in an attempt to advance their own interests. One’s sense of a moral identity seems to be key [19]. It seems that a sense of power can activate one’s personal traits in a way that can have an impact on their moral identity. Perhaps, too, power can be expressed differently depending on moral identity. To have a moral identity means that an “individual holds morality as part of his or her self-concept” [19, 31]. The extent to which one has a moral identity also seems to dictate one’s emphasis on their own versus another’s needs [19, 31]. With one’s moral identity, comes the question of how they act when placed in a position of power. In a study by DeCelles, power was defined as a psychological state in which one perceives a high level of control, which in turn leads to particular cognitive and affective changes as well as tendencies towards certain actions [19]. Recall the 5 bases of power that were previously outlined by French and Raven. Incorporating those definitions as well as the concept of evil outlined above, one can begin to see publication bias unfolding.

Ultimately, in this study by DeCelles, it was shown that “power predicts self-interested behavior differently depending on moral identity.” So, when someone has a low moral identify and is provided power, they are more likely to act in their self-interest as they find morality to be unimportant to their self of self or character. The opposite is true for those with a high moral identity and power, as they find morality and their moral identity to be central to their sense of self. This is a factor when it comes to decisions about publishing studies. While moral identity is a hard trait to decipher from a job interview, or through training, we as a community should strive to foster environments in which morality is constantly checked, strengthened, and where there are high expectations to always act with the common good at the forefront. If we fail to do so, we may find ourselves in a situation in which the implications of publication bias are great and cause irreparable harm – incomplete and inadequate information provided to further appropriate disease-specific research, incomplete guidelines to keep the population healthy and safe, and a mistrust of the scientific institution.

An additional argument that must be made when it comes to the ethical obligations one has in research deals with the participants, or human subjects. When enrolling in a clinical trial or whatever the study may be, human subjects are doing so not to gain anything individually, but to hopefully better the lives of others in the future and add to the development of scientific breakthroughs and understanding. This is a basic tenet of enrolling in and giving informed consent for a study one plans to participate in. Since these participants knowingly exposed themselves to some level of risk, known or unknown, we should be morally obliged to publish all results [6]. Not only do we have an obligation to the greater good, but these individual obligations are important as well. In good faith, we must continue this trusting relationship with research participants if we hope to continue to conduct meaningful studies with willing participants.

III: FEMINIST & VIRTUE ETHICS

Feminism Overview

Feminism strives to attain justice for women, and feminist ethics approaches philosophical problems in a similar manner. Feminist ethics, in the sense that it will be used to discuss publication bias, is a theory that values inclusivity, moral emotions, narratives, and the importance of the lived human experience [32, 33]. Feminist ethics recognizes that the traditional philosophical theories often overlook vital aspects of an argument. Feminist ethics, for the most part, separates itself from traditional theories and claims that any sort of bias is detrimental to a fair and logical analysis of a situation or problem. It challenges the idea that philosophy has always been about supporting humanity and striving for the betterment of it.

Feminist ethics pushes to uncover the biases within each traditional moral theory, specifically bias toward the experience of men – that the male experience just was moral experience. Recognizing the flaws and trying to right the wrongs done to those not supported previously is a core tenet of feminist ethics. In general, feminist ethics argues that every person, no matter their social status or place in a social hierarchy, should have moral agency and be free to operate in a world that respects this [34]. Additionally, some feminists contend that individualism is a masculine value and idea which negatively impacts one’s sound reasoning. Traditionally, ethics of care has been contrasted with ethics of justice, but it has been said to be a dangerous dichotomy to draw as it undermines the view of many feminist scholars [35]. This is relevant here, as made evident in the previous chapter, due to the strong power dynamics involved in the process of research. This chapter will explore how feminist ethics would view the problem of publication bias and how the interest and importance of the whole, rather than the particular individual, must be considered when researching, writing, editing, and publishing.

Bias, Objectivity, and Empiric Research

“The human sciences do not constitute a reservoir of factual knowledge, uncontaminated by values…” [36]. This statement by Alison Jaggar in her book *Feminist Politics and Human Nature* provides an appealing introduction to the incorporation of science into philosophy. Often thought of as two separate entities that do not agree on much, science and philosophy both provide insights on publication bias. The imperfections of human nature will impact the process of science, the interpretation of science, and sometimes even the validity of science. Additionally, every human has a biased value system shaped and refined based on each individual’s experiences. This creates a lens through which everything else is seen, including science [37]. Because feminist ethics values a person’s lived experience to inform the ethical framework, there are certain expectations of a researcher that should be enforced and fostered. In Brabeck’s The Handbook of Social Research Ethics, in which a feminist perspective is utilized to outline some certain responsibilities, several questions that a researcher should ask are laid out [35]:

1. “Whose voices are left out of the research?
2. What populations are ignored in the study of the phenomenon under investigation?

3. Which experiences are not given scholarly attention?”

These questions are relevant when assessing published research and ought to encourage research that is both inclusive and non-biased. When the literature consists of a majority of studies with only significant and positive results, it is clear a subsect of that research is quite simply missing and perhaps purposefully hidden.

A cornerstone of science is that of objectivity. To introduce subjective reasoning and analyses into science would threaten the respect and importance it holds. But when we consider the ethics of publication bias, we must recognize that it is nearly impossible to be completely objective in research. In every field, including that of research, the values a society holds are influential and create biases. The motivation for a topic that is being researched and the topic itself can be influenced by society’s values, as our cultures are engrained in each one of us [38]. Ethical views differ tremendously, but in practical and realistic contexts, we have to negotiate and accommodate. The influence of societal values is important to recognize when contemplating how processes (i.e. publishing studies) can be improved. This already introduces some level of subjectivity. To be truly objective, the needs of the society would not be considered. However, this is not possible. Rather, objectivity is about the gradual, critical understanding garnered from several overlapping and individual analyses. This view should lend itself well to science, as the formation of respected scientific fact is based in exactly that – repetitive critical examination. While scientists and those involved in publishing research ultimately regard objectivity highly, they are pulled in several directions – is this significant? Is it relevant? Is it going to gain traction in an environment that is inundated with new findings and advancements? Additionally, from the very beginning of the research process, the theories and hypotheses posed were likely never fully objective themselves. They were molded by what the society around it values and what external factors were driving it [38]. Further, as discussed earlier, the body of existing literature and knowledge greatly impacts what is prioritized for further research and publication. The publication of a negative study would be important in that it would indicate a path that may not be relevant for future research. It would be important to publish this so that others do not waste time and money on asking the same questions, and achieving the same results. If published, other researchers could determine an alternate question or approach, which could further medical progress. This is another example of how objectivity can be further undermined if publication bias continues. To further understand and determine how objectivity can be upheld in science, philosopher Hugh Lacey's dichotomy can be utilized.

Lacey defined two different versions of objectivity – that of neutrality and impartiality.

> "Neutrality is the requirement that scientific theories neither presuppose nor support any non-epistemic (moral, political, social, or cultural) values. Impartiality requires that theories be evaluated on the basis of evidence and the extent to which they exhibit other epistemic values." [38]

When contemplating these two complementary versions in the context of publication bias and research in general, it seems impartiality can help us in determining steps to decrease the frequency of publication bias as well as recognizing when we may be stepping into risky areas. Being impartial does not mean the interests and values of society and individuals are ignored, but rather, to have the ability to look past individual values and recognize the importance of research with the greater good in mind, rather than that of a specific entity's personal goals and interests. One must recognize when the biases held, or the background information informing one's decisions, are beneficial or harmful. Simply being aware of these influencing factors can help increase the level of objectivity and ensure a fairer and more just way forward. This version of objectivity relates nicely to empiricism in feminist theory, as outlined by philosopher of science Helen Longino. Longino argues in her work that it is impractical to expect science to be value-less. It is the context and background values, rather, that make it meaningful and impactful. The same values and interests that may be harmful to the process, may in fact lead to constructive criticism and the improvement of the process as a whole. It is again important to balance this with the constant recognition that it is present and influential.

Virtue Ethics

Oftentimes, moral theories are seen as too reductive, making it somewhat impossible to apply concepts to real-life cases [39]. In light of this, a normative approach can be taken as well. Virtue ethics emerged in a time when people were dissatisfied with Kantian and Consequentialist theories when discussing right action and virtue [39]. While Kantian approaches focus on a set of defined principles in order to live and Consequentialist theories focus on the results or consequences of an action, both allow for virtues to be taken into account, but virtues are not foundational to their theories. Virtue ethics puts virtues front and center. Within this theory, virtues are what guide actions and from which other acts will be rooted within [40]. In virtue ethics, virtues are not simply traits that one does or does not possess, they are fundamental to living and living well. Virtue ethics mainly focuses on the type of person one aspires to be. The moral questions about what qualities one should have to be morally upstanding are dependent on virtues, thus virtue ethics lends itself to the topic at hand.

But what exactly is a *virtue*? And what does it mean to be *virtuous*? First, we must distinguish between a *virtue* and a *moral virtue.* Beauchamp and Childress explain that a *virtue* is “a dispositional trait of character that is socially valuable and reliably present in a person, and a *moral virtue* is a dispositional trait of character that is morally valuable and reliably present” [40]. Therefore, a moral virtue does not rely on societal norms or values and should, then, be consistent throughout different cultures, religions, and nationalities. Moral virtues should theoretically be aspirations of the many and be held to the highest of expectations. A moral virtue can also be thought of as “the disposition to act or a habit of acting in accordance with, and with the aim of following, moral principles, obligations, or ideals” [40]. Simply put, virtues are excellences, which exist over a wide array of activities, and moral virtues are habits that manifest good or bad human action. But this also means that every moral virtue is derived from foundational beliefs, many of which are relative and dependent on the society in which it is formed, which comes from some place of value.

In order to understand this further, we can consider one’s intentions or motives. These, in turn, impact their character and provide us with something to judge and value. When motivated by morally virtuous characteristics (such as sympathy), we approve and could call that person virtuous. When one is motivated by self-serving motives, though, we do not prescribe the characterization of virtuous to them [40]. We therefore deem it necessary for the motive behind actions to be in good nature as well, meaning with morally virtuous intentions and actions. Without that key component, one’s actions can be as good, but cannot be classified as *virtuous* without a proper motive [40].

Studies have continuously shown that many studies are not published, reportedly as many as 50% [41]. The negative impacts of selective, biased publishing are extensive. Not only can they be damaging to the progress of science, but there can actually be real consequences on the health of patients. After all, “decisions made about patient care, protection from hazards, and lifestyle recommendations are made based on consideration of the whole literature, not just a single study” and “a clinical treatment may be considered effective when reviewing literature that is subjected to publication bias, but this consideration can be found to be erroneous when all evidence is considered” [41]. If we were to be truly virtuous, we would take a preventive approach to this and ensure that the possibility of such a biased scope of literature does not come to be in the first place. We see the process of writing, publishing, and researching as an intellectual practice (or set of practices), so we can appeal to a standard of excellence for that practice (i.e. what makes for the successful completion of the task) and define the virtues as those habits that enable us to achieve the goods of the practice. Therefore, the virtues of responsible research would, from this standpoint, be wary of and prevent publication bias.

Aristotle's doctrine of the mean[3] is also useful in these kinds of situations. While it is hard to really pinpoint what it means to attain human Eudaimonia[4], it seems realistic "that a moral virtue is a mean – "an intermediate condition" – between two extremes" [39]. In order to make the decision in the moral conflict of 1) publishing only those with the most significant positive results and 2) publishing everything that is submitted, one must reach a middle ground, or rather, the mean. The mean, and ideal outcome as described by Aristotle and also within this thesis, would be publishing all well designed studies whether or not they are positive results. Courage is one virtue that can be applied. Someone who possesses the virtue of courage has to have a balance of fear and cowardliness [39]. We could argue that pursuing option 1 could stem from an extreme, such as fear of losing traction in one's field or fear of harsh reviews for publishing something some do not expect. It could also be argued that pursuing option 2 could stem from another extreme, fear of retaliation if one were to reject too many studies, or fear of choosing the "wrong" studies to publish. Finding a balance that allows

---

[3] Aristotle's doctrine of the mean, simply put, is the idea that a *virtue* is the mean, or average, of two extremes. It is the happy place; in which everything is found to be in balance. [39]

[4] Eudaimonia is a state of happiness or flourishing, first defined by Aristotle. It is often used as a grounding framework for virtues. [39]

for the virtue at play to take over would lend itself to the best overall outcome, and a just distribution of published studies and results. Therefore, exhibiting the virtues enables one to reach a more just publication process.

The Nature of Flourishing

The meaning of flourishing and the path to it has been a long contested topic, but it is one that in this context, and many others, extends the feminist and virtue ethics argument. Ultimately, we hope research is assisting in the journey to flourishing, but for whom and how do we achieve this? What does it mean to truly flourish? Who defines that? Is it dependent on an individual's values or is it more inclusive than that? Does flourishing simply pertain to a single person at a time, or does it require an entire society in order to be attained? Just about every philosopher has a different answer. Additionally, the larger and broader topic of human nature must be taken into account. Under the umbrella of human nature, what comprises "human abilities, needs, wants and purposes" must be considered [36]. These are all highly contested concepts themselves, as certain needs and wants can be counted as Jaggar puts it, "real," while others are purely perceived as necessary by the individual. When a closer look is taken, it becomes clear that how these are defined brings the central argument down to the concept of flourishing. Further, "what constitutes flourishing and well-being, however, is clearly a question of value, both with respect to the individual and with respect to the social group" [36]. Simply put, in order to flourish, one must be morally virtuous in every facet of their life. This not only includes acting in the best interest of themselves, but in the best interest of others. Taking into account and acting in accordance with the welfare of others is crucial when it comes to human flourishing.

How does human flourishing pertain to the issue at hand, publication bias? Each and every action is connected and plays a part in the success or failures of the consequential result, so when researchers prioritize, or rather de-prioritize, certain research, the downstream consequences must be considered. These downstream consequences, I will argue, pertain to the general public who relies on sound medical knowledge to inform their treatment teams when they face an illness. The medical community has certain norms, values, and expectations. Those values are centered around a common goal: to place the patient first. And so, research for this community should abide by the same values and mission. Placing the health and safety of those who will benefit from the information gained from research should be the top priority. Without health, an integral component of well-being, additional values cannot be achieved and thus do not promote flourishing of any kind. At this point, public research can be declared a common good that can only be achieved as a part of an intellectual effort on the part of researchers who have certain professional and social virtues. If the achievement of this common good (responsible results of research) are necessary for a flourishing medical community, then we have to design research practices and guidelines that support it and inculcate virtues that can sustain it. For those involved in conducting and publishing research, it seems the morally right thing to do would be to take the necessary steps to aid in the flourishing of the field, and thus the flourishing of those who are impacted by such findings.

IV: A UTILITARIAN APPROACH

In order to argue that researchers and medical journals ought to produce unbiased and honest work, a consequentialist moral theory can be utilized. Classic utilitarianism is arguably the most common form of consequentialism and it has largely been formed thanks to the influence of Jeremy Bentham and John Stuart Mill. Utilitarianism, at its core, focuses on the consequences of an action and derives moral value from consequences rather than whether the action was necessarily right or wrong [39]. The action *is* right or wrong to the extent that it produces greater consequences (greater overall balance of pleasure over pain, for example). Utilitarians also stress pursuing actions that result in the most happiness and flourishing for the population; it argues for the greater good [39].

This approach to solving moral conflicts is quite similar to a public health approach, which is why it lends itself well to the scenario at hand. Part of the foundation of utilitarianism, the Greatest Happiness Principle, explains that by creating an ideal that is attained by working towards the greatest happiness for the greatest number, a morally acceptable outcome can be achieved [42]. In order to understand this more fully, we have to consider the long term impacts of publication bias, many of which have already been discussed. Some of these include distortion of data, the negative impact on the perception of the integrity of research, decreased trust in the literature being used to make medical decisions, the impact this may have on patients if the data is incomplete, and mistrust in the system.

Mill's Greatest Happiness Principle, otherwise known as the Utility Principle, explains that in order to reach the ultimate state of being, as a whole, we must maximize happiness of each individual [43]. The sum of everyone's individual happiness is what allows us to reach this ideal moral society, which Mill says is what defines what is moral or not [43]. In thinking about this point, it is also important to recognize that there may be instances in which one person pursues an action that makes them happy, but in turn causes harm to another, thus increasing the happiness of one but decreasing the happiness of another. In this case, where one person's pleasure is equivalent to another person's pain, there is no utility produced – a zero sum action. This does not necessarily mean we should not pursue that action, as ultimately the goal is to pursue the action or set of actions that result in the most happiness for the majority, even if that means decreasing the happiness of some. In moving forward with this, it would allow (and require) the appropriate publication practices even if it meant negatively impacting some involved (journal's impact factor, editor and publisher's salaries, etc.) – after all, utilitarianism allows one to sacrifice the few for the good of many.

However, this this line of thinking exposes a flaw or complication in the Greatest Happiness Principle. We could see this play out in a situation where the overarching research field created a policy in which 100% appropriate publication would occur, allowing for the greater medical community to have the complete data available in order to make the appropriate decisions. But in doing so, the journals may suffer in terms of decreased impact factors and therefore a lower reputation leading to decreased revenue which would in turn negatively impact those who work for them. So how do we come to terms with these conflicting experiences? This may be where the Harm Principle plays a role. The Harm Principle works to minimize pain and suffering, in whatever form it may arise, by prohibiting one entity from causing harm to another [43]. In the case of publications, this could be utilized in order to allow (and require) the appropriate publication practices without negatively impacting the livelihoods of those who work for the company. Whether or not the publication industry could be regulated is debatable, however it would be possible to improve the culture surrounding it, the expectation for the scope of studies journals should be publishing, and perhaps the way we “reward” journals with impact factor scores. The culture of research would have shifted due to these principles, thus leading to an environment in which all involved are the happiest they can be, living and working in the most moral society possible. This supports the decision to require and strongly advocate for a culture shift in the research world, so that publication bias becomes a problem of the past. Otherwise, it remains a threat to the greater good. For utilitarians, the ultimate value is in that of increasing utility – in this case increasing the worth and inclusiveness of medical literature to benefit the most people – which can be achieved if the goal of publication shifts to be the most informative and complete body of work.

An additional component of utilitarianism is to choose the action that reduces suffering, as alluded to by the Greatest Happiness Principle and Harm Principle. In a world focused on alleviating the suffering of others, it seems counterintuitive to allow for a certain practice to continue when it has been shown to distort the data that is presented, as previously mentioned [14-16]. The immediate impact on one’s understanding of a disease process, treatment, etc is not the only worrisome factor, the long-term and downstream effects are paramount. In addition to the immediate and short-term impacts it may have on a body of literature, an incomplete compilation of research may drastically skew what continues to receive funding, the scope of a project, and what continues to be published. This could lead to the suffering of many, if they are not receiving the best possible care due to biased research practices, as well as suffering for those conducting research if there are continued biased publishing practices. Unlike hedonists, utilitarians believe that the conglomeration of everyone's happiness, over individual happiness, is the most important and therefore, the morally right action is one that results in the least amount of negative utility (or bad consequences) (in this case, missing information) and the most amount of pleasure.

There are two different kinds of utilitarianism: act utilitarianism and rule utilitarianism. If operating within the act utilitarianism arena, one runs the risk of harming a person if the action of doing so leads to the greater good for everyone else [42]. This proves treacherous when applying the theory to certain situations. Due to this, rule utilitarianism says that we should rather adopt societal rules and principles if and only if they produce the greatest good for the entire society [42]. In applying rule utilitarianism to publication bias, we can argue that the rules selected for research purposes can be shown to produce better results (unbiased reports). However, the problem with rule utilitarianism is that there are some circumstances in which violating the rules results in more utility, and if that is so, the rule is undermined.

By creating expectations that negative results are equally as likely to be published, given the research methods and writing is sound, the larger validity of the literature would be intact and strengthened. By creating a space in which it becomes routine to submit studies that do not reach significance for review and publication, the stakes (for reputation within the scientific field, for popularity in who submits studies for review, for individual career goals, etc) will not be as high and a much more well-rounded experience would be achievable. For if the main outcome and goals are centered on what each study would add to the literature and overall well-being of those impacted, rather than an individual entity (whether that be a company or person), the entire system would be focused upon the greater good. This would value both the lives of an entire population as well as the lives of individuals.

The utilitarian argument does not take into account individual rights [42]. Autonomy is highly valued in society, so many may ask how we can morally override an individual's right to choose what they do or do not choose to publish (or even pursue). While a valid concern, it has been shown that transparency in research promotes scientific progress, accountability, and replicability [44-46] and could in fact maximize good. This lends itself well to the common good, a vitally important goal for utilitarians. The common good, for utilitarians, is an aggregate concept; individual rights and pleasures do not matter. Therefore, the safety and wellbeing of the public, through trustworthy research and publications, should be deemed more important than a single person's feeling of autonomy. In addition, one may argue that creating an environment in which this occurs consistently, the rate of publication bias would decrease and the validity of the literature would increase.

In the end, health is a common, shared, and public concern. If we continue to allow for individuals and groups with agendas to make these important decisions, we are complicit in putting society at risk. We also walk a fine line next to a slippery slope – one where groups can decide what research and which researchers are worthy of publication. Skewing information like this does more harm than good and cannot continue to be the norm.

V: A KANTIAN APPROACH

When we discuss autonomy as lay persons, we usually mean the freedom of a person to pursue whatever desire they feel inclined to pursue. This definition is quite loose and permits a wide variety of actions, some good and some bad. Philosophers are not so general though and oftentimes discuss autonomy in three different ways – personal autonomy, moral autonomy, and political autonomy. Moral autonomy as discussed by Kant, is defined as the “capacity to deliberate and to give oneself the moral law, rather than merely heeding the injunctions of others,” while personal and political autonomy do not involve morals, they rather focus on the capacity of one to pursue whatever one wishes to and to have those decisions respected [47].

Kant established moral autonomy due to his dissatisfaction with allowing outside forces and governing bodies to dictate what principles we use to make decisions. He emphasized an individual’s ability to self-govern and stressed the importance of following one’s own law of morality [47]. Kant also excluded feelings, emotions, and all other non-intellectual factors from his account of autonomy, leaving a decision-making route to be fact-based, logical, and rational. In Kant’s world, moral agents are ends in themselves – one develops and abides by moral laws in a society where others are developing and abiding by their own moral laws, which may or may not be different from one another [47]. So if each person acts individually and follows a diverse set of individual moral laws, then how are we to have a moral world? Kant thought that the maxims we are guided by may be able to be generalizable, thus worthy of being obeyed and respected by all rational beings [39]. The concept of autonomy is crucial here. Humans, as rational and autonomous agents, are capable of setting expectations for themselves and creating rules that they would follow. This is central to the Categorical Imperative. In order to fulfill the Categorical Imperative, one must have autonomous will [48]. The first law, within the Categorical Imperative, dealt with universalizability (or universal law) – *“act only according to that maxim by which you can at the same time will that it should become a universal law”* [49]. In other words, only act on policies and principles that could be reasonably endorsed by all who have to live under them. These are important notions to understand when determining the right course forward in publication, and other fields. Perhaps a reframing of our ultimate goals in publishing research (improve the health and safety of our society, provide honest and complete data on all topics, etc.), following the Categorical Imperative could be helpful in addressing change.

In Kant’s philosophical endeavors, he sought to provide a method to answer the question, “what ought I do?” The Highest Good, the quintessential goal of human life, was also noteworthy to Kant and he said that in order to achieve it, complete moral virtue must come together with complete happiness [48]. He did not believe this to be a simple quest, and postulated that we all have moral duties that must play a role in answering that question. These duties include duties to ourselves as well as duties to others, encompassing a whole host of moral obligations. These duties and values were to be categorized, and have often been said to be “metaphysical” and “unachievable,” but Kant was convinced that they were actually incredibly rational and provided an overriding reason for certain actions [48]. All of the abovementioned, but especially the duty to others, is what makes it possible that Kant would have found it imperative to establish and enforce an objective, fair, and just process for publishing research findings. Relevant here is Kant’s distinction between “perfect” and “imperfect” duties – perfect duties being those you are not permitted to refrain from and imperfect duties being those you act upon for personal fulfillment (i.e. charity and generosity) [39]. This leads me to the conclusion that there may very well be perfect duties that relate to fair research practices. A classic example of a perfect duty is to tell the truth, and so we should never lie to uphold that perfect duty. In publication, this could be applied to honesty and transparency in which studies are and are not published as well as the reasoning behind such decisions. If there is a perfect duty to publish the “right” papers (i.e. those that have sound methodologies and investigate an important question that will further the field), then we should not dismiss or reject them based on significance of results or level of interest the review board holds in the study.

In this chapter, when autonomy is referred to, we will be discussing the autonomy of those at the receiving end of research - meaning, the medical providers who utilize the information and data that is published to guide their practices, the governing medical bodies that create guidelines and develop protocols based on the evidence, and finally the patients who are autonomous when it comes to deciding on a course of treatment and other medical decisions. These are each individual moral agents. Shouldn’t we protect the ability for these others who are not directly involved in research to be autonomous? To be truly autonomous requires appropriate and accurate information that allows for one to have the best understanding of their individual situation. Without research that portrays this, we are not enabling our patients and medical teams to make well informed autonomous decisions.

Whenever autonomy or decision making capacity are discussed in medicine and in ethics, there is always a vital component that cannot be ignored and that is the information the patient is presented with. The information should not be presented in a biased manner, nor should it be incomplete or misleading. If the information, at large, in the medical community is incomplete due to the problem of publication bias, we cannot be reliably certain that the information our patients are utilizing to make decisions is the best it could possibly be.

Kant would have also addressed intention. In deontological theories, intention is important. For Kant, one’s primary motives and intentions have to be committed to morality, and the virtues are those strengths of character that make it possible to do that [39]. So if researchers, editors, and publishers are keeping this information with the intention of improving or safeguarding the scientific community, it may be justified. However, it is difficult to imagine a situation in which that would be the case at all times. As discussed previously, Kant’s value of duty is paramount. In the medical research community, ethical standards have been decided upon and can be interpreted as one’s duty in that role. The Code of Medical Ethics, set forth by the American Medical Association, has published the following as a guide for physicians involved in biomedical and health research, specifically speaking to the obligations that they have when it comes to research dissemination with the “ultimate benefit” of physicians and patients in mind:

- “Advocate for timely and transparent dissemination of research data and findings. Physicians should not intentionally withhold information for reasons of personal gain.
- Report the results of research accurately, including subsequent negative findings. This is particularly important where the findings do not support the research hypothesis.
- Maintain a commitment to peer review.
- Disclose sponsorship and conflicts of interest relating to the research, in keeping with ethics guidance.
- Be responsible in their release of research results to the media, ensuring that any information the researcher provides is prompt and accurate and that informed consent to the release of information has been obtained from research participants (or participants’ legally authorized representative when the participant lacks decision-making capacity) prior to releasing any identifiable information.” [50]

The AMA guidance indicates that research findings should only rarely be withheld and if withheld, “then only to the extent required to reasonably protect against misuse” [50]. When contributing to such a large body of information and to a rapidly changing and evolving field, the utmost respect for transparency and honesty should be at the forefront. While this Code of Medical Ethics exists, it is unclear how it is implemented and enforced, other than relying on the good will and citizenship of researchers. Also, if we as research entities are making decisions that will impact other people down the line, should we not always consider the best interest of those people who will be impacted? In doing so, we are being the most virtuous versions of ourselves. Kant believes that virtues are “strengths of will” that enable someone to obey the moral law and commit oneself fully to morality [39]. For Kant, morality regards others, meaning that in order to act morally one must act for the sake of law, not for the sake of self-interest. Just as is expected in the clinical world, researchers, editors and publishers should place the needs of others ahead of their own self-centered gains and value the importance of those individuals who will be impacted. If studies are continuously filed away or rejected due to their “uninteresting” or negative results, we are lying by omission.

VI: COUNTER-ARGUMENTS

While the purpose of this thesis was to explore the ethical implications of publication bias, as well as the presence of it, it is prudent to address the fact that there may be ethically sound reasons that lead to a majority of positive studies being published. It would be foolish to assume that all negative studies that are not published are due to ill intentions. There are many reasons why certain studies are not submitted for review or published and it is important to outline those.

For one, not all studies are of high quality. Some studies may have resulted in negative results purely because of being poorly defined including an inaccurate hypothesis. Studies that are poorly designed would not add to the literature and should not published. Studies with a small sample size, inappropriate statistical analysis, outdated results given other published studies since the start of the study, and the absence of consistent methodology are some examples that could lead to non-publication, other than bias [51]. These rejections can aid researchers in further refining their study to be one that is carried out appropriately. One must remember that most published studies, which this study analyzed, are a small subset of all studies and are theoretically the best of the best given the rigorous review process.

Additionally, researchers may “file drawer” studies due to realistic and appropriate reasons – lack of time being the biggest reason. Some do not have the time to publish every study and may be using certain results to guide future studies which will be submitted for publication. At the researcher’s level, decisions are made on what they deem will be the most useful to the field [51]. If this is not done excessively and is done without the pressures of outside entities, it is ethically sound and can be seen as indeed a good motive.

Ultimately, not everything can be published. That is an impossibility. Editors and reviewers have to make decisions, many of which are subjective. Every journals provides a guide to how each manuscript is evaluated, the so called “Journey of a Paper” on the Lancet’s website, and argues in favor of selection criteria [52]. Decisions must be made based on themes of the issue and novelty of the results, amongst others [51]. This automatically leaves many studies to be rejected, simply based on the interests of the journal or the field’s popular area at the time. However, this does not mean biased selection is justified. We must continue to work towards non-biased selection and review, extending our ethical standards to publication.

VII: METHODS

Specific research questions

This thesis sets out to examine publication bias in medical journals, both recently and over the years. There are two ways to measure publication bias: indirectly and directly. In order to indirectly measure publication bias, one can assess whether positive findings (defined as statistically significant results supporting the stated hypothesis) in published studies and larger effect sizes in smaller studies are disproportionately overrepresented in the literature [3]. Previous studies have also described more direct measures such as comparing published and unpublished studies, and studies that followed up cohorts of studies [3]. However, the problem with the former direct measure is that it is usually impossible to know how many studies remain unpublished, especially if they are not clinical trials [3]. In this project, studies will not be chosen based on whether or not they have follow up studies available for comparison. Every study will be included, regardless of follow up status. Since the studies will be different and examine unrelated hypotheses they cannot be compared to one another, my analysis will simply examine the proportion of positive studies among all studies in a pre-determined sample of journals. The information gathered from this will allow me to draw conclusions about submission and publication practices. The main questions that will be examined include:

1. What proportion of published hypothesis-testing studies reported positive results (i.e. observed results that support one or more of the proposed hypotheses)
2. How does this proportion differ across types of journals and study types?
3. Has this proportion changed over time?

I hypothesize that there will be a higher proportion of positive results throughout, that there will be more negative results published as the years progress due to increased attention and knowledge of publication bias as well as the requirement to register clinical trials.

Data sources and selection of journals and years of publication (process and rationale)

Medical journals were first organized by impact factor (IF), listed highest to lowest. Journals were excluded if they were:

- not written in English,
- limited to a specific country or specialty (to maintain focus on broad scope international journals),
- did not have articles published in 1 of the 3 years under investigation (1998, 2008, and 2018)

Individual articles were excluded if they presented:

- case reports,
- systemic reviews and meta analyses,
- descriptive analyses without addressing a testable hypothesis

The highest 3 journals by impact factor that resulted included the New England Journal of Medicine (NEJM), the Lancet, and the Journal of the American Medical Association (JAMA). The lowest 3 journals by impact factor included the Journal for International Medical Research, Military Medicine, and Internal Medicine. Impact factors were as follows – NEJM (74.7), the Lancet (60.4), JAMA (45.5), the Journal for International Medical Research (1.3), Military Medicine (1.03), and Internal Medicine (1.01).

Impact factor is a common measure utilized to rank the “importance” of a journal, with higher impact factors being more desired and respected. Impact factors are calculated based on how often they are cited and utilized over a two-year period by “dividing the number of times articles were cited by the number of articles that are citable” [53]. Ranking by impact factor is a contentious topic as many argue that it is in and of itself a biased and inaccurate measure; one that is self-perpetuating and does not necessarily mean high quality research [54]. Nevertheless, this is the most common method for determining the hierarchy of journals and does lend itself well to this study.

In order to analyze any time trends that may be present, three years each ten years apart were determined – 1998, 2008, and 2018. Of note, in 2007, it became required that all clinical trials be registered. By choosing this time span, we will be able to analyze any impact this may have had on the proportion of positive versus negative studies published. Additionally, this time span appeared to be a reasonable amount of time to determine any trends given the increased attention and knowledge surrounding publication bias.

Criteria used to select articles for the final analysis

Within in the high impact factor group of journals (NEJM, the Lancet, and JAMA), all original, hypothesis-testing studies from all issues in each designated year were included.

In order to compile a sample of lower impact factor journals, a random 25% of issues from each of the three designated years were chosen and all original, hypothesis-testing studies within those issues were compiled. A sample, rather than all, was utilized due to time constraints. In the future, we may examine all of the low impact journal articles. An online random number generator - calculator.net - was used to identify which issues would be used from each journal and year. The random selection was as follows:

- Journal of International Medical Research - 2018: issues 1, 2, and 10; 2008: issues 1, 4, and 6; 1998: issues 3, 5, and 6
- Military Medicine - 2018: Issues 1/2, 3/3; 2008: issues 4, 7, and 11; 1998: 1, 9, and 10
- Internal Medicine - 2018: issues 2, 4, 11, 18, 20, and 24; 2008: issues 4, 6, 8, 16, 18, and 22; 1998: issues 2, 4, and 8

Since each article was required to be original and hypothesis-testing, this excluded case reports, briefs, perspective pieces, systemic reviews, meta-analyses, and descriptive studies. Once this was determined, each article’s title, first author, and DOI were entered into an excel sheet where the rest of the coding and final analysis would take place.

Specific information extracted from each article that met inclusion criteria

For each article that met inclusion criteria, two main pieces of information were extracted: study type and result category (positive statistically significant vs. not statistically significant). For the purposes of this study, there were two main study types – experimental and observational. Observational studies include both analytical and descriptive, but since descriptive studies were excluded this just left analytical, which we will refer to as observational in nature. In the excel database, there were coded by letters – E for experimental and O for observational. In basic terms, experimental studies are those in which an investigator assigns exposures. Observational studies are those in which exposures are not assigned. In order to differentiate between analytical and descriptive, one must determine whether or not there was a comparison group (figure 1 and 2). Usually the study type could be determined from the abstract, specifically the methods section. Rarely, it was unclear, and a full text review was required.

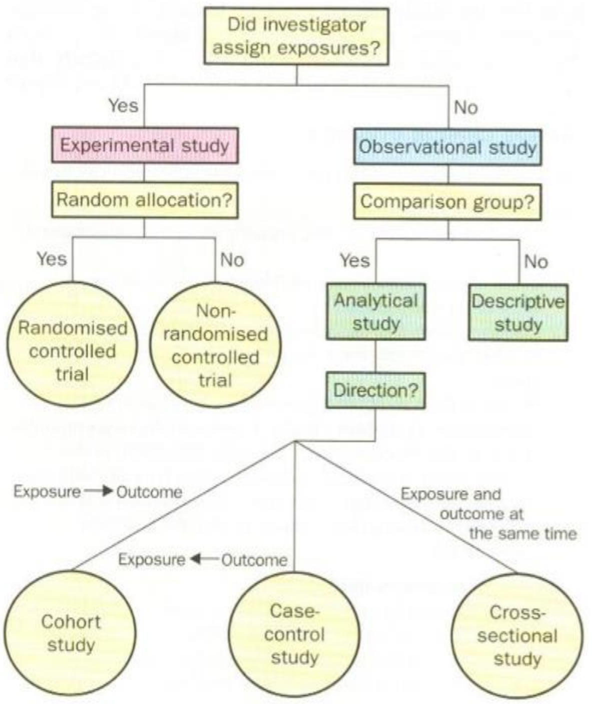

Figure 1: Algorithm for classification of clinical research [55].

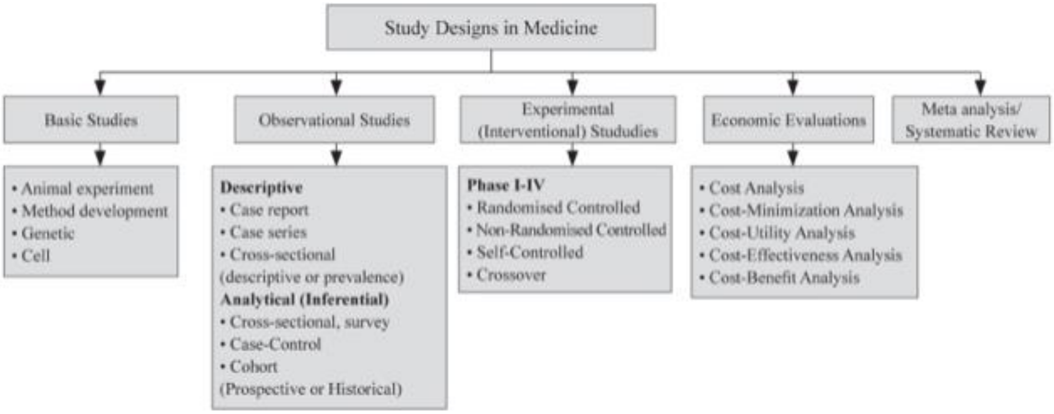

Figure 2: Study Designs [56]

Once the study type was determined, the result was assigned into one of two categories: positive statistically significant (denoted as ‘positive’) or not statistically significant (denoted as ‘null’). Only primary outcomes were considered. A positive significant result was defined as rejecting the null hypothesis (Ha) and a non-significant (‘null’) result was defined as failing to reject the null hypothesis (Ho). For trials, specifically, it was necessary to first determine whether it was a superiority, non-inferiority, or equivalence trial, as this defined the null hypothesis. For reference:

- Superiority trial hypotheses:
  - Ho: There is no statistically significant difference
  - Ha: There is a statistically significant difference
- Non-inferiority trial hypotheses:
  - Ho: There is a statistically significant difference
  - Ha: There is no statistically significant difference
- Equivalence trial hypotheses:
  - Ho: They are not the same
  - Ha: They are the same

Once the category of results was determined, it was either given a 1 or a 0 in the database. A 1 indicated a positive result and a 0 indicated a null result.

Second Coder

Of note, a second coder was utilized to code a random 10% of the articles. This served as a way to validate the original coding method and allowed for adjustments to coding methods, if necessary. A random number generator - calculator.net - was used to determine 10% of the line numbers in the database, which correlated to an individual article. The second coder was then provided a blank excel document with only those journal articles listed and coded study type and significance as they saw fit.

Following the second coding, both coders re-examined any disagreements and came to a final decision. If no final decision could be agreed upon, the study in question was to be sent to the thesis committee for review. This was done in two waves – 10% of the high impact factor journals (NEJM, the Lancet, and JAMA) was done first. Then, 10% of the low impact factor journals (Journal of International Medical Research, Military Medicine, and Internal Medicine) was done.

Second Coder Results

Following exclusion of descriptive studies, the sample selected for re-coding included 150 studies (experimental or observational). Among those, there was a 91% agreement. The coders disagreed on 14 studies. 2 of the disagreements were study type (E vs. O), and the remaining 12 disagreements were the type of results (positive vs. null). After review, an agreement was reached on all 14 of the original disagreements. 11 of 14 were changed to the 2nd coder's result and 3 of the 14 remained as had originally been coded. Therefore, 11 of the 150 (7%) did not match the original coding. There was ultimately about a 93% accuracy. Notably, the majority of the disagreements (9/14) came from NEJM - which is the journal both coders started with. Both agreed that coding became easier the more one did and that NEJM was the journal with the most confusing language.

The second 10% re-coded 37 studies. There was a total of 84% agreement. The coders disagreed on 6 studies. 5 out of the 6 disagreements were on study type, while 1 out of the 6 disagreements was on type of results (positive vs. null). After review, 100% agreement was reached. 3 out of 6 were changed to the 2nd coder's result and the rest remained as had originally been coded. Therefore, 3 out of 37 (87%) did not match the original coding. This resulted in about a 92% accuracy amongst the low impact factor journals

Data analysis

All analyses were conducted using the Statistical Analysis System (SAS) version 9.4 statistical software. Descriptive analysis calculated frequencies and percentages for all variables. A bivariate analysis was conducted to describe and compare the outcomes across categories based on journal title, journal group and study type. Changes in the proportion of positive results were examined over time both overall and within each journal and article category. Statistical inferences were carried out using chi square tests for comparison across categories and Cochrane-Armitage tests for trend to assess for changes over time. Two sided p-values of <0.05 were used to define statistical significance.

VIII: RESULTS

Publication bias in medical journals can prove to be a detriment, with many ethical implications. This study set out to indirectly measure publication bias in 6 medical journals at three definitive time points. The questions that were analyzed, again, are as follows:

1. What proportion of published hypothesis-testing studies reported positive results (i.e. observed results that support one or more of the proposed hypotheses)
2. How does this proportion differ across types of journals and study types?
3. Has this proportion changed over time?

Study sample (Table 1)

The total sample size in this study was 1773 articles. Of those, 1429 (81%) reported positive results while 344 (19%) reported null results (Table 1). There were 6 total journals analyzed – the New England Journal of Medicine (NEJM), the Lancet, the Journal of the American Medical Association (JAMA), the Journal of International Medical Research, Military Medicine, and Internal Medicine. NEJM had 569 articles (32%), the Lancet had 432 articles (24%), and JAMA had 479 articles (27%) (Table 1). The sample of lower tier journals resulted in the Journal of International Medical Research having 152 articles (9%), Military Medicine having 76 articles (4%), and Internal Medicine having 65 articles (4%) (Table 1). Sorted by year, there were 605 articles (34%) published in 1998, 574 articles (32%) published in 2008, and 594 articles (34%) published in 2018. When broken down by study type, 877 articles (49%) were classified as observational while 896 articles (51%) were experimental (Table 1).

Table 1: Descriptive Characteristics of Study Sample

	n	%
Positive Results	1429	81
Null Results	344	19
Journal		
The New England Journal of Medicine (NEJM)	569	32
The Lancet	432	24
The Journal of the American Medical Association (JAMA)	479	27
Journal of International Medical Research	152	9
Military Medicine	76	4
Internal Medicine	65	4
Year		
1998	605	34
2008	574	32
2018	594	34
Study Type		
Observational	877	49
Experimental	896	51

Association between journal, journal group, and study type (Table 2)

Journal title

When each journal was considered individually, the NEJM published 461 (81%) positive studies, the Lancet published 355 (82%) positive studies, JAMA published 368 (77%) positive studies, the Journal of International Medical Research published 127 (84%) positive studies, Military Medicine published 61 (80%) positive studies, and Internal Medicine published 57 (88%) positive studies. The variation across these proportions was not statistically significant (p = 0.1532).

Journal Group: High vs. Low Impact Factors

The high impact journals published 1184 (80%) positive studies, whereas journals with a low impact factor published 245 (84%) positive studies. This difference was also not statistically significant (p = 0.1525).

Study Type

Is there an association between study type and the likelihood of reporting a positive result? Among 897 experimental studies, 652 (73%) reported positive results. The corresponding proportion was significantly higher (89%) among 876 observational studies. This difference was statistically significant (p<0.0001).

Table 2: Bivariate Analyses of Journal, Journal Group, and Study Type with Study Significance

	Positive		**Null**		**X^2 P value**
	n	**%**	**n**	**%**	
Journal					
NEJM	461	81	108	19	0.1532
Lancet	355	82	77	18	
JAMA	368	77	111	23	
Journal of International Medical Research	127	84	25	17	
Military Medicine	61	80	15	20	
Internal Medicine	57	88	8	12	
Journal Group					
High Impact Factor Group	1184	80	296	20	0.1525
Low Impact Factor Group	245	84	48	16	
Study Type					
Observational	778	89	99	11	<0.0001
Experimental	651	73	245	27	

Changes over time (Table 3)

Overall

Overall, a significant downward trend was observed (p = 0.0004). Whereas articles published in 1998 included 85% positive studies, this proportion decreased to 77% by 2018.

By Impact Factor

The decrease in the proportion of positive studies was especially pronounced among high impact journals, (p<0.0001). In 1998, 479 (86%) positive studies were published. In 2008, 356 (78%) null studies were published. In 2018, 349 (75%) positive studies were published.

The corresponding time trend was no longer evident in the group comprised of low impact journals, (p = 0.5450). Overtime, there was no significant change in the number of positive or null studies that were published. In 1998, 37 (80%) positive studies were published. In 2008, 98 (84%) positive studies were published. In 2018, 110 (85%) positive studies were published.

By Study Type

Within observational studies, the proportion of positive results decreased significantly over time (p= 0.0281). In 1998, 313 (92%) positive observational studies were published. In 2008, 266 (87%) positive observational studies were published. In 2018, 199 (86%) positive observational studies were published.

By contrast, the same trend was less discernable and no longer statistically significant (p = 0.1772) when the analysis was limited to experimental studies. In 1998, 203 (77%) positive experimental studies were published. In 2008, 188 (70%) positive experimental studies were published. In 2018, 260 (72%) positive experimental studies were published.

Table 3: Examination of temporal changes in the proportion of positive results by journal group and study type: 2008-2018

Journal Characteristics	1998				2008				2018				Test for trend P-value
	Positive		Null		Positive		Null		Positive		Null		
	n	%	n	%	n	%	n	%	n	%	n	%	
By Impact factor													
High	479	86	80	14	356	78	101	22	349	75	115	25	<0.0001
Low	37	80	9	20	98	84	19	16	110	85	20	15	0.5450
By Study Type													
Observational	313	92	28	8	266	87	39	13	199	86	32	14	0.0281
Experimental	203	77	61	23	188	70	81	30	260	72	103	28	0.1772
Overall	516	85	89	15	454	79	120	21	459	77	135	23	0.0004

IX: DISCUSSION

Interpretation of Findings

This study revealed that amongst the six journals analyzed, 80.6**%** of published hypothesis-testing studies reported positive results. There was no significant association found between journals in their likelihood to publish positive studies. However, a low tier journal, Internal Medicine, was the journal that published the highest proportion of positive studies (87.7%). Between journal groups (high vs. low impact factor), there was no significant association in proportion of positive studies published. Between study types, there was a significant association. Null studies were more likely to be published if they were experimental studies rather than observational studies. Several factors could be the reason for this. First, trials are required to be registered which may lead to an incentive to submit all research for publication. Additionally, there is more money in experimental studies than in observational studies which could also drive researchers to aim for publication no matter the result. Because observational studies are not as well funded and do not require as much money to be carried out, unimpressive results may be filed away. This is not to say that the results are unimportant as they likely help to guide subsequent studies.

Over time, a significant change in proportion of positive studies published was found. From 1998 to 2018, the proportion of null studies published significantly increased. If broken down by impact factor, significantly more null studies were published in high impact journals as time progressed (1998 to 2018). In 1998, only 14.3% of studies in the high impact journal group were null studies, but in 2018, 24.8% of published studies were null studies. However, within the low impact group of journals, there was no significant change over time. These findings may also be due to the fact that in 2007, a subset of experimental studies (trials) were required to be registered which may have increased the incentive to publish results. Additionally, as the years have progressed, there has been increased attention and knowledge surrounding publication bias which may have led to journal review boards as well as researchers being more inclined to submit and publish null studies, as was discussed in depth in the introductory chapter. Simple awareness potentially impacts one's biases and allows for different decisions to be made. If broken down by study type, a significant change over time was found within observational studies. Over time, significantly more null observational studies were published, with 8.2% published in 1998 and 13.9% published in 2018. Again, this is likely due to increased awareness and attention on the problem that is publication bias [3]. There was no significant change over time for experimental studies, with 23.1% published in 1998 and 28.4% published in 2018. This lack of change over time could be explained by the fact that experimental studies have always required more funding, buy-in, and effort [57]. Thus, I hypothesize that they seem to automatically be more likely to be published, or at least submitted, even when the results do not confirm the hypothesis.

Ultimately, since it is difficult to directly quantify publication bias, this indirect method is useful and has brought to light many important points. From a purely proportion standpoint, more positive results were published than null results. It could be said that there are disproportionately frequent positive findings in published studies. Further, both significant and non-significant associations were confirmed between different study types, journals, as well as over time.

Notable Observations

There were several notable observations made while compiling data. In several journals it was noted that when one study found no significant results, another study with the same authors in the same journal would immediately follow. This subsequent study would pose a different research question that would ultimately result in a significant or positive outcome. This was interesting as it allowed for the null study to be published, as well as a new positive outcome from the same study. It would be worthwhile, in the future, to investigate whether or not this 2nd paper and 2nd research question was delineated prior to the outcome of the 1st study or if it was created following the initial outcome(s).

Additionally, many studies had two primary outcomes defined and would find one to be positive (significant) and the other to be null (non-significant). The way this data was compiled, this result would be counted as a positive result. In the eyes of an editor, this would also be seen as a positive study. We must rely on honest and ethical research design in this case. If both outcomes were primary prior to the results being determined, this would be a legitimately positive study. However, if there was just one primary outcome at the start of the study and it was found to be non-significant, a significant secondary outcome may have been re-designated a primary outcome. This could lead to the proportion of studies being skewed towards those that are positive when in fact, they would have initially been negative.

Another important observation while collecting data was that some journals were clearer than others in terms of a hypothesis. Many articles did not explicitly state their hypothesis, but rather left the reader to infer or assume what it was intended to be. Could this make the study more likely to be accepted for publication even if there was a negative result? Of note, JAMA seemed to be the most well-organized of the journals in this review in terms of studies clearly stating hypotheses and results.

Finally, in the lower tier journals, there were many more pre-clinical experimental studies. These were often carried out in mice, so there may be less pressure and money invested which may lead to any outcomes being more likely to be published. In addition to this, the fact that they were in the lower tier journals seems to confirm the hypothesis that lower tier journals are less selective and may be influenced less by the medical and research society to publish positive studies.

Limitations

Some limitations of this study include the potentially imperfect coding system. While a set method was used to determine what kind of study each was, it is possible there were some errors made. Additionally, the lower tier journals only consisted of a sample of 25% of studies. If we were to collect data from all studies in these three journals, the data could be more complete. However, due to time limitations and the purpose of the study, it was not deemed necessary.

X. CONCLUSION & REFLECTIONS

This thesis has confirmed and challenged many of my previously held thoughts surrounding medical literature and the publication bias that may or may not be present, as well as some of the potential implications of its presence. Counter-arguments enlightened my understanding of the topic as well, and allowed for an honest analysis of the data in that I had to recognize that null studies are not always worthy of publication. This challenged my bias that I may have entered this study with – that top tier medical journals are “too good” for null studies and are not doing their part to thoroughly review the studies submitted to them. It is unfair to say this, as the review process often taken several weeks to months and reviewers are highly trained and knowledgeable persons. Recognizing that they in fact are human, and may make mistakes, is necessary. It does not excuse biased methods, but does add a layer of complexity that cannot be ignored.

Perhaps the most important overarching takeaway regards that of the intentions of researchers and those involved in publishing studies, as well as the environment in which this all takes place. This is important when it comes to trusting the greater system of research (from researchers to reviewers to publishers) to ethically determine what studies are published. Those involved in publishing studies should be involved for the right reasons, broadly being the betterment of others and in pursuit of the common good. With these intentions, we would all be off to a better start when it comes to the problem of publication bias. In order for this to occur, it lies in the culture of the environment in which this work takes place. In other words, it depends on if and how the culture of an institution (or more largely, a society) fosters moral identity. Previously, I proposed that we as a community should strive to foster environments in which morality is constantly checked and strengthened, and where there are always high expectations to act with the common good as the priority. This is relevant not only in the publication bias realm, but in almost every facet of our world. How do we achieve this? Interestingly, many have deliberated on the importance of this and how it can be carried out in practice. In Monaci's essay on embedding humanizing cultures in organizations, two major approaches are taken to describe one's moral agency – particularly as an organization or community [58]. First, Monaci describes the Catholic social doctrine's view on this topic as one that views the enterprise "as a community of work, which leads to place stress on the possibility of creating 'organizational humanizing cultures' that revolve around the principles of human dignity and the common good and allow organization members to flourish" [58]. Further, the Catholic social doctrine emphasizes that because every individual who works in a group is capable of forming bonds with others, this forces the often self-interested motives to dissipate or become questioned [58]. With this comes the increasingly natural proclivity to work towards shared goals that go beyond themselves as well as beyond the organization. This also allows for greater meaning to be obtained from the work and therefore lead to increased flourishing for both the workers and those who may benefit from the work being done. In this case, that could include patients and medical providers.

The second point that Monaci makes relates to how this culture can be instilled by leadership by utilizing Philip Selznick's, a sociologist, viewpoint. In summary, Selznick proposes that the role of leadership is critical and can instill a sense of moral agency via ascertaining one's moral identity in every individual who works for them [58]. With strong and supportive leadership, a culture that values the good their work can do can be attained. By having leaders who exhibit moral actions and are able to recognize immoral actions and act when they witness these occurring, an environment will be propagated that will propel us to a higher level. By having a role model in a leadership position, we can begin to instill moral thought and aspirations in every other person involved. Additionally, this serves as support for the aspiration that all participants in research and publication should be working towards the betterment of others and in pursuit of the common good. With intentions that are set in this, we can expect publication bias to become less and less frequent.

Recent Events

Recently, there has been attention brought to the publication of studies as it relates to the COVID-19 (coronavirus) pandemic and it brings to light issues that are relevant to the topic at hand, especially the integrity and honesty involved in reviewing studies prior to them being accepted or rejected for publication. In early 2020, as the rapid emergence of SARS-CoV-2 gained international attention, medical journals were publishing studies rapidly. A new disease that was spreading fast rightly generated scientists to do what they do best – learn about it and share that information via papers in medical journals. However, as studies were quickly published, several were also quickly redacted due to flaws in their claims and less than ideal methodologies. These papers that were published all had one thing in common – stunning and seemingly game-changing information. Or, rather, as described in this thesis, positive results. One study “promised that popular blood-pressure drugs were safe for people infected with the coronavirus" while another study stated that hydroxychloroquine, a medication thought it in the early stages to help with coronavirus infections, was actually harmful to patients [59].

Once they were published, medical professionals and others involved immediately criticized them declaring fraudulent claims and deception. They were quickly redacted after these claims came to light and were verified. Backlash centered primarily on the mistrust that studies like these could cause, and warned against the pressures reviewers for medical journals face [59]. These studies were not found in "low tier" journals, either. They were accepted and published in the top medical journals – The Lancet and The New England Journal of Medicine (NEJM), which is perhaps even more troubling. Were they just publishing to be the first to have groundbreaking research on a new and frightening pandemic, or were the peer reviewers rushed in an unprecedented time? Either way, if they were not held accountable by outside forces, the implications could have been great. Who is to say this has not happened several other times, but just has not been caught?

It appears that some of the studies that were redacted during this time period were actually fraudulent in that data was manipulated and made to appear perfect, when in fact the data did not exist. If authors are to be trusted, this cannot be tolerated. The process relies on humility and honesty, so we cannot allow anyone to participate in this important field of work if they are to be influenced by morally wrong goals. The peer-review process is supposed to protect against these mistakes, and are vital when it comes to transparency in research. Simply publishing studies which show interesting results is no way to practice. There must be a balance in times like these between speed (which is necessary, as time is of the essence when dealing with a rapidly spreading infectious disease) and honestly critical review. The lives of people are quite literally at stake, and we should be demanding the most stringent of processes in order to produce the most trustworthy and unbiased research.

References:

1. Joober, R., et al., *Publication bias: what are the challenges and can they be overcome?* Journal of psychiatry & neuroscience : JPN, 2012. **37**(3): p. 149-152.
2. Dickersin, K. and Y.I. Min, *Publication bias: the problem that won't go away.* Ann N Y Acad Sci, 1993. **703**: p. 135-46; discussion 146-8.
3. Hooper, S. and Y. Loke, *Publication bias: what is it? How do we measure it? How do we avoid it?* Journal of Clinical Trials, 2013. **5**: p. 71-81.
4. Guan, M. and J. Vandekerckhove, *A Bayesian approach to mitigation of publication bias.* Psychon Bull Rev, 2016. **23**(1): p. 74-86.
5. Easterbrook, P.J., et al., *Publication bias in clinical research.* Lancet, 1991. **337**(8746): p. 867-72.
6. Mlinarić, A., M. Horvat, and V. Šupak Smolčić, *Dealing with the positive publication bias: Why you should really publish your negative results.* Biochemia medica, 2017. **27**(3): p. 030201-030201.
7. Duyx, B., et al., *Scientific citations favor positive results: a systematic review and meta-analysis.* J Clin Epidemiol, 2017. **88**: p. 92-101.
8. Fanelli, D., *Negative results are disappearing from most disciplines and countries.* Scientometrics, 2012. **90**(3): p. 891-904.
9. Fanelli, D., *Do pressures to publish increase scientists' bias? An empirical support from US States Data.* PLoS One, 2010. **5**(4): p. e10271.
10. Dwan, K., et al., *Systematic review of the empirical evidence of study publication bias and outcome reporting bias.* PLoS One, 2008. **3**(8): p. e3081.
11. Murtaugh, P.A., *JOURNAL QUALITY, EFFECT SIZE, AND PUBLICATION BIAS IN META-ANALYSIS.* Ecology, 2002. **83**(4): p. 1162-1166.
12. Stern, J.M. and R.J. Simes, *Publication bias: evidence of delayed publication in a cohort study of clinical research projects.* BMJ, 1997. **315**(7109): p. 640.
13. Ekmekci, P.E., *An increasing problem in publication ethics: Publication bias and editors' role in avoiding it.* Med Health Care Philos, 2017. **20**(2): p. 171-178.
14. Devito, N. and B. Goldacre, *Publication Bias.* Catalogue of Bias, 2019.
15. Goldacre, B., et al., *Compliance with requirement to report results on the EU Clinical Trials Register: cohort study and web resource.* BMJ, 2018. **362**: p. k3218.
16. FDAAA. *Who's sharing their clinical trial results?* 2019 [cited 2020 September 28]; Available from: https://fdaaa.trialstracker.net/.
17. Lakens, D. and A.J. Etz, *Too True to be Bad: When Sets of Studies With Significant and Nonsignificant Findings Are Probably True.* Social Psychological and Personality Science, 2017. **8**(8): p. 875-881.
18. Dickersin, K., *How important is publication bias? A synthesis of available data.* AIDS Educ Prev, 1997. **9**(1 Suppl): p. 15-21.
19. DeCelles, K.A., et al., *Does power corrupt or enable? When and why power facilitates self-interested behavior.* J Appl Psychol, 2012. **97**(3): p. 681-9.
20. Weber, M., *Economy and Society: An Outline of Interpretive Sociology*, ed. E.F.e. al. 1978, Berkeley, CA: University of California Press.
21. Hobbes, T., *Leviathan*. 1985, New York: Penguin Books.

22. French, J. and B. Raven, *The bases of social power*. Vol. 6. 1959.
23. Marcovitch, H., *Editors, Publishers, Impact Factors, and Reprint Income.* PLOS Medicine, 2010. **7**(10): p. e1000355.
24. Largent, E., et al., *Misconceptions about coercion and undue influence: reflections on the views of IRB members.* Bioethics, 2013. **27**(9): p. 500-507.
25. Morgenthau, H.J., *The Evil of Power.* The Review of Metaphysics, 1950. **3**(4): p. 507-517.
26. Collins, P. *Power and Choice: A Philosophical Inquiry into Evil*. 2013 [cited 2021 Janaury 25]; Available from: https://www.utsa.edu/ovations/vol8/story/power.html.
27. Cadler, T., *The Concept of Evil*. 2018, Stanford Encyclopedia of Philosophy.
28. Calder, T. *The Concept of Evil*. 2020 [cited 2020 December 13]; Available from: https://plato.stanford.edu/archives/sum2020/entries/concept-evil/.
29. Hwang, T.J., et al., *Failure of Investigational Drugs in Late-Stage Clinical Development and Publication of Trial Results.* JAMA Internal Medicine, 2016. **176**(12): p. 1826-1833.
30. Lee, K., P. Bacchetti, and I. Sim, *Publication of clinical trials supporting successful new drug applications: a literature analysis.* PLoS medicine, 2008. **5**(9): p. e191-e191.
31. Aquino, K. and A. Reed, *The self-importance moral identity.* Journal of personality and social psychology, 2003. **83**: p. 1423-40.
32. McAfee, N. *Feminist Philosophy*. 2018 [cited 2020 September 30].
33. Gotlib, A. *Feminist Ethics*. [cited 2020 September 30]; Available from: https://iep.utm.edu/fem-e-n/ - H2.
34. Garry, A. *Analytic Feminism*. 2012 [cited 2020 October 3].
35. Brabeck, M.M. and K.M. Brabeck, *The Handbook of Social Research Ethics*, ed. D.M. Mertens and P.E. Ginsberg. 2013: SAGE.
36. Jaggar, A.M., *Feminist Politics and Human Nature*. 1983, Totowa, New Jersey: Rowman & Allanheld.
37. Kuhn, T., *The Structure of Scientific Revolutions*. 1962: University of Chicago Press.
38. Crasnow, S. *Feminist Perspectives on Science*. 2020 [cited 2020 December 8]; Available from: https://plato.stanford.edu/entries/feminist-science/ - Obje.
39. Timmons, M., *Moral theory : an introduction.* 2013.
40. Beauchamp, T.L. and J.F. Childress, *Principles of Biomedical Ethics*. Seventh ed. 2013, New York, New York: Oxford University Press.
41. Shields, P.G., *Publication Bias Is a Scientific Problem with Adverse Ethical Outcomes: The Case for a Section for Null Results.* Cancer Epidemiology Biomarkers &amp; Prevention, 2000. **9**(8): p. 771.
42. *The Nature of Morality and Moral Theories*. [cited 2019 November 22]; Available from: http://home.sandiego.edu/~baber/gender/MoralTheories.html.
43. Nahra, C., *The harm principle and the greatest happiness principle: the missing link.* Kriterion: Revista de Filosofia, 2014. **55**: p. 99-110.
44. Knottnerus, J.A. and P. Tugwell, *Promoting transparency of research and data needs much more attention.* J Clin Epidemiol, 2016. **70**: p. 1-3.
45. Prager, E.M., et al., *Improving transparency and scientific rigor in academic publishing.* Brain and behavior, 2019. **9**(1): p. e01141-e01141.
46. Selvan, M.S., et al., *Ethics of transparency in research reports.* Indian J Med Ethics, 2011. **8**(1): p. 31-6.

47. Dryden, J., *Autonomy*, in *Internet Encyclopedia of Philosophy*. 2019.
48. Johnson, R. and A. Cureton, *Kant's Moral Philosophy*, in *Stanford Encyclopedia of Philosophy*. 2016.
49. Driver, J., *Ethics: The Fundamental.*, ed. Blackwell. 2007.
50. AMA. *Principles for Disseminating Research Results*. 2020 [cited 2020 October 8]; Available from: https://www.ama-assn.org/delivering-care/ethics/principles-disseminating-research-results.
51. Thrower, P. *Eight reasons I rejected your article*. 2012 [cited 2021 March 21]; Available from: https://www.elsevier.com/connect/8-reasons-i-rejected-your-article.
52. Lancet. *Journey of a Paper*. 2020 [cited 2021 March 22]; Available from: https://www.thelancet.com/journey-paper.
53. *Measuring Your Impact: Impact Factor, Citation Analysis, and other Metrics: Journal Impact Factor (IF)*. 2021; Available from: https://researchguides.uic.edu/if/impact.
54. Seglen, P.O., *Why the impact factor of journals should not be used for evaluating research.* BMJ (Clinical research ed.), 1997. **314**(7079): p. 498-502.
55. Grimes, D.A. and K.F. Schulz, *An overview of clinical research: the lay of the land.* The Lancet, 2002. **359**(9300): p. 57-61.
56. Süt, N., *Study designs in medicine.* Balkan medical journal, 2014. **31**(4): p. 273-277.
57. Health, I.f.W. *Observational vs. experimental studies*. 2016 Available from: https://www.iwh.on.ca/what-researchers-mean-by/observational-vs-experimental-studies.
58. Monaci, M., *Embedding Humanizing Cultures in Organizations through 'Institutional' Leadership: the Role of HRM.* Humanistic Management Journal, 2020. **5**(1): p. 59-83.
59. Rabin, R.C., *The Pandemic Claims New Victims: Prestigious Medical Journals*, in *The New York Times*. 2020.