

Through the One-Sided Mirror: Effects of Commodification on an Individual's Cultural Perception

Muxin Yao^{1*}

¹Phillips Academy, Andover, MA USA

Received September 18, 2022; Revised November 26, 2022; Accepted, December 7, 2022

Abstract

Earlier literature and research have examined cross-cultural differences in cognition and behavior across a multitude of fields, spanning sociology, business, and psychology. The scope of such research on cultures has mainly focused on describing the culture, or the individual as part of that culture; in comparison, there has been little investigation of the effects of cultural interactions. Research thus far has investigated subjects such as consumer cultural reactance, as well as responses to cultural identity threats. However, despite this literature, and research on the disparity between internal and external perceptions of self, little literature exists regarding how one views their cultures. Foreign audiences may see a culture different from the people who identify with that culture. This disparity may be further emphasized during cultural branding and marketing when only limited aspects of that culture may be emphasized for monetization. This paper aims to explore how these differences in cultural and self-perception affect one's perception of their own culture and one's behavior with respect to that change after the culture has undergone commodification and marketing to a foreign audience. We focus on the nuance within this and how the behaviors may change depending on whether the culture is individualist or collectivist, the level of identification to a culture, and the salience of monetary profit from commodification. Through this, we hope to utilize the literature on cultures, on self vs. other perceptions, and on the commodification of in-groups vs. out-groups to develop a framework for assessing identifier reactions to cultural commodification and intercultural interaction.

Keywords: Cultural Commodification, Cultural Perception, Individualism vs. Collectivism, Self-Identification, Cultural Disparity

1. Introduction

Globalization has greatly increased the reach of originally domestic products and services, allowing the culture they represent to be transported across the world in the luggage within commercial planes and crates aboard cargo ships. The rapid growth in intercultural interaction is inspiring many countries to promote their own culture, seeking to improve their reputation to both gain profit and elevate their worldly status (Subramanian, 2017; Zax, 2013).

The Cool Japan Strategy is Japan's attempt at

country branding. Since the end of World War II, Japan has been focusing its cultural diplomacy on the softening of anti-Japanese receptions, particularly in Southeast Asia (Iwabuchi, 2015). Historically, Japanese politicians and academics have viewed their culture as being unique in comparison to the rest of Asia (Moeran, 1983). Now, Japanese cultural industries are continuing the self-perceived distinction politicians and academics have had, seeking to become the "interpreters of 'the West'" for Asia (Iwabuchi, 1998). With those ambitions in mind, Cool Japan presents three main objectives: promoting

* Corresponding Author
muxin.yao@gmail.com

Advisor: Andrew Morningstar
atssmith@wharton.upenn.edu

growth in the Japanese population's understanding of western culture, increasing the presence of Japanese perspectives around the world, and rebranding Japan as a country that "helps the world" (Cool Japan Proposal, 2014). For example, the plan proposes increasing English education and holding international festivals in Japan to respectively improve the Japanese people's understanding of western culture and attract the attention of foreign tourists, increasing Japan's global presence (Cool Japan Proposal, 2014). By improving individual communication through language education and inviting tourists to visit Japan through festivals, this strategy aims to cultivate collective change from the individual level.

It is important to acknowledge the significance authenticity has in judging cultures, given the impact that cultural marketing has on authenticity (Newman, 2014). The consumption of cultural products is motivated by people's desire to obtain an authentic cultural emblem or experience, such as people traveling abroad to try regional cuisines rather than eating a familiar dish at home (Newman, 2014). For a country that is trying to commercialize its culture to outsiders, the aspects they commercialize may depend on what is demanded by foreign consumers. As a result, the representation of that culture to the foreign audience may be different from what is represented to the home culture, questioning whether the commercialized version of a culture may be perceived as authentic by the home culture. Intentional branding of cultures also might demean their products' authenticity as people may question whether the original culture has undergone modification to increase its appeal. Altogether, authenticity is a powerful determinant of consumer preference and aversion (Suter, 2014; Silver, et al., 2020).

There is a need to study the effects that these differences in perception have on people of the home culture— including changes in the culture and their relationship with the culture—to balance cultural preservation and cultural exchange amidst commodification and globalization. This research presents a novel perspective because there exist no papers on how an outsider's perception of a culture may be different from an identifier's perception of

that same culture. There are also no studies on how that mismatch affects identifiers of the perceived culture. Yet, cultural interactions and resulting changes prove to be a common phenomenon as cultures constantly brand and reinvent themselves to outsiders for purposes of tourism and exports (Cool Japan Proposal, 2014; Theodoropoulou and Tovar, 2021). To develop an understanding of the discrepancies between a culture's self-perception and how that culture is being represented, and the resultant changes in cognition and behavior, reviews were made from several separate literatures that are distinct from but applicable to this topic, including identifiers' associative and dissociative response to cultural brands, and assimilative and reactive behavior when exposed to foreign cultures (Hornsey and Jetten, 2005; Jetten and McAuliffe, 2002; White and Argo, 2009; White, et al., 2012; Kim, et al., 2014). By bringing together this series of extended literatures, this paper is able to provide a framework for understanding the effects of cultural commodification. It is hoped that this paper spurs further research in line with the propositions made and finds balance amidst the cacophony of potential inauthenticity.

2. Existing Literature

Numerous papers have been published investigating the independent effects of individualist vs collectivist cultures, high vs. low identifiers, and their interactive effects with the presence of monetary benefit. As it is impossible to present all of the existing research, this review examines a subset of those literatures relevant to their effect on cultural self-perception and behavior. In particular, papers were selected for their analysis of the behavioral characteristics of individualists and collectivists, as well as case studies examining the reactions of individuals towards the commodification of a culture, whether that be of their own or of a culture they are interested in purchasing the commodity (e.g., through tourism) for. These papers are then cross-referenced with each other to formulate conclusions regarding the various trends observed within. Rather than generally examining the effect marketing has on cultures, the review seeks to uncover the ways in

which the commodification of cultures impacts the ways identifiers and outsiders view that culture, and therefore how that changes the home culture and the people who belong to that culture. With that, the marketing choices made about cultures are merely a byproduct of cultural perception, and what is considered to be of the value of those cultures by outsiders.

This paper will illustrate the concepts and propositions primarily using Japanese culture. Japan is a country that is heavily commodified with a high number of cultural exports such as anime and video games. It also receives a lot of attention in international tourism (Uppink and Soshkin, 2022). This paper also utilizes case studies from Japanese culture and Japanese reactions to commodification as illustrations of its conclusions. These illustrations will be buttressed in places with examples from other cultures.

2.1 Individualist vs. Collectivist Cultures

Of the literature investigating individualist vs. collectivist cultures, the two most important subliterations for understanding the capitalistic export of culture are the means by which an individualist or collectivist understands the self and how an individualist or collectivist understands in-groups and out-groups in relation to the self.

Understanding the Self

The differences in which individualists and collectivists understand themselves can be based on two dimensions: independence and self-consistency (Hornsey and Jetten, 2005; Kim, et al., 2014; Cousins, 1989). Studies indicate independence to be viewed as a desirable trait by individualists (Hornsey and Jetten, 2005). The more important a factor is in establishing or enhancing self-worth, the more people will emphasize their person around that factor (Vazire and Carlson, 2011). Such is the case of independence for individualists, who have demonstrated such a desire for independence that they will rationalize away their conformist behavior (Hornsey and Jetten, 2005). This may also be explained by the individualist's high degree of self-consistency, more specifically that they will erase their conformist

behavior to maintain a consistent individualist persona they hold for themselves (Hornsey and Jetten, 2005). Also as a result of self-consistency, those individualist cultures remain unlikely to align their self-view to what they perceive others' evaluations of them to be when they do not match (Kim, et al., 2014). The characteristic to remain true to their self-perception makes the mindsets and behaviors of individualists less affected by situational cues, demonstrating the overall cultural emphasis on self-consistency and psychological autonomy (Kim, et al., 2014). In contrast, the collectivist person's self-perception is much more fluid. When asked to describe themselves, identifiers of collectivist culture use less abstract values (i.e. generalized character traits) and instead rely on their social role, for example, a woman identifying as a mother in one situation while being a wife in another. These social role identities may change with context, reflecting lower self-consistency (Cousins, 1989). This "non-western concreteness"—the tendency to focus on situational behavior and social role—is attributed to the collectivist's cognitive inability to summarize inconsistencies in behavior and is seen as a way to keep perceiving the individual as being dependent on context and group norms (Cousins, 1989).

Means By Which People Understand Others and the Group in Relation to Self

As aforementioned, individualists assign value to themselves based on their degree of independence from the group (Hornsey and Jetten, 2005). As such, they view group influence as potentially compromising individual potential and therefore will prioritize personal goals over collective goals (Hornsey and Jetten, 2005). Not only will they reject group influence, but individualists are also biased towards egocentric projection, the projection of their own emotions onto others (Cohen and Gunz, 2002). Referencing a psychological perspective on the situations, it means that individualists prefer taking a perception of the world where they use existing information to understand new information (Cohen and Gunz, 2002). Contrastingly, collectivists then opt for adjusting their known information around the new information and utilizing more relational projection where they project onto others emotions that they

stereotype the other to feel rather than assuming the other to feel the same as themselves (Cohen and Gunz, 2002). To better illustrate this phenomenon, consider a situation where an individualist or collectivist is conversing with someone they just met and wants to evaluate how the stranger feels about a certain topic. The individualist may begin with the assumption that the stranger feels the same way about the topic as they do, while the collectivist will be devoid of that assumption and instead observe the stranger's actions to make their inference on how they expect the stranger to feel. As a result, the individualist subconsciously prioritizes the self in all logical and emotional evaluations, while the collectivist prioritizes the group.

2.2 High vs. Low Identifiers

According to Social Identity Theory, a group is defined as “a collection of people who categorize themselves as belonging to the same social category and internalize the category's social identity-defining attributes to define and evaluate themselves—attributes that capture and accentuate intragroup similarities and intergroup differences” (Tajfel and Turner, 1986). Within that group, individuals may display a range of identification—the strength of their association and similarity between their personal identity and that group's identity. Previous research has identified that those with different levels of identification react differently to group-related stimuli (ScienceDirect). Currently, research on identification has focused on mainly two categories: exhibited behavior when encountering identity threat, and level of conformity to group norms.

Constructs Related to Identification

In their 2009 paper, White and Argo introduce the term “collective self-esteem” or “CSE,” defined as “the degree to which the individual sees the self as a worthy member of the social group, values the social identity, and sees the social identity as important to the self-concept.” By that definition, having a high CSE is similar to being a high identifier of a culture and vice versa. Both CSE and high vs. low identification exist as valid metrics of an individual's

connection to their group.

Another similar concept research has addressed involves the concept of independent and interdependent self-construals (White, et al., 2012). At its core, self-construal refers to “whether individuals view themselves as primarily separate from or integrally connected to others.” (Giacomin and Jordan, 2017). However, self-construals are a more general measure of a person's degree of social reliance rather than on a particular culture. For example, one can be independent or interdependent with their friends, but may not describe their friend group as individualist or collectivist. As someone with a more interdependent self-construal is more likely to subscribe and connect with a group, those with interdependent self-construals are more likely to be high identifiers of a culture and vice-versa.

Behavior in Response to Identity Threat

Social identity threat occurs when an aspect of a person's social identity is devalued or in danger of disappearing (White and Argo, 2009). After examining existing literature, responses to identity threats relevant to cultural interactions can also be examined from two different angles – the associative or dissociative behavior one will demonstrate to their own culture, and the negativity of attitudes displayed towards the foreign culture causing the threat.

Both an individual's tendency to display associative vs. dissociative response and their attitude towards the invading culture relies on the strength of their attachment to a culture, whether that may be the home culture or the foreign culture. In both the studies conducted on CSE and self-construals, people with high CSE or an interdependent self-construal prefer identity-linked products even when that identity is threatened; conversely, those with low CSE or independent self-construals will avoid products related to the threatened identity (White and Argo, 2009; White, et al., 2012). These choices are explained by each group's approach to restoring self-esteem after facing identity threat: those with independent self-construals do so through dissociating from the threat while those with interdependent self-construals find a sense of belonging through connecting with the group identity (White, et al., 2012). The principle of identification

can also be applied to investigating one's attitude towards an invading culture by setting the invading culture as the identity threat in those aforementioned situations.

Another area research has examined involves reactions to identity threat through collective identity closure—the negative attitude one demonstrates towards a foreign group or perceived identity threat and the resultant closing-off of that group to out-groups (Morris, et al., 2011). The closure is usually done to narrow the definition of the in-group and exclude minorities (Morris, et al., 2011). Exposures to situations where the home culture is mixed with a foreign culture will evoke the need for closure due to identity threat being triggered through distinctiveness loss and assimilation (Morris, et al., 2011). However, this behavior is more observed in those with a low identification with the foreign/invading culture and less in those with a higher identification (Morris, et al., 2011). By extension, those with more independent self-construals or lower CSE will likely be less inclined to display closure when threatened.

Conformity to Group Norms.

When discussing conformity, the majority of existing research does so within the premise of individualism vs. collectivism. A group norm of individualism will encourage higher identifiers to act for the benefit of the self and collectivism for the benefit of the group (Jetten, et al., 2002). For collectivists, their stances will be relatively straightforward, with high identifiers adhering to the group norm and lower identifiers less so (Jetten, et al., 2002). However, a dilemma presents itself for high identifiers of individualism as their inclination to stand by their group makes them collectivist by definition, but simultaneously contradicts the individualist norm (Jetten, et al., 2002). As such, high identifiers of individualist cultures may purposely show individualism as a form of conformity (Jetten, et al., 2002).

2.3 Presence of Monetary Benefit

Monetary benefit is a crucial component in the discussion of cultural commodification as

commodification is primarily driven by the desire for profit, making it significant to examine the impact profit has on cultural goods and services. Existing literature mainly explores the management of cultural odor—the presence of the producer's cultures—within the goods themselves, as well as tourism's impact on the advertised culture's local communities.

Cultural Odor

Some degree of cultural odor—the presence of the producer's culture—is present in every product, and firms are able to intensify or diminish this through intentional marketing (Iwabuchi, 1998). The term *mu-kokuseki* originates from the Japanese term describing something that lacks a nationality, and in the marketing context refers to products that are relatively culturally odorless (Iwabuchi, 1998). For example, while the burger chain McDonald's began as an American brand, its cultural ownership has become insignificant in the consumer's eyes after its stores opened across the world and developed menu items based on regional dishes (Iwabuchi, 1998). However, a product designed with a lack of intentional cultural odor may still contribute towards the consumer's perception of a culture's odor—a product may carry cultural odor even if it was deliberately designed to lack odor as the medium through which the product is delivered may carry odor that the producer themselves are not aware of. This phenomenon is particularly emphasized in Japanese animation. For the most part, characters and settings in Japanese animation often deliberately do not look ethnically Japanese, the art style choosing to erase particular characteristics and cultural context (Iwabuchi, 1998). However, western audiences, after watching these animations, will associate the odorless, *mu-kokuseki* style to be from Japan and thus redefine their perception of Japanese cultural odor to be derived from the portrayal of an animated, virtual Japan devoid of bodily ethnic characteristics (Iwabuchi, 1998).

While cultural odor is not necessarily diminished through international commodification, it can be altered, partially by localizing the changes to match the demands of foreign audiences of a particular background. The changes that occur can be within the product's purpose and usage—such as McDonald's

developing region-specific menu items—to the product’s presentation, such as changing the language on the packaging or the appearance of cultural symbols (e.g., flags) to match the one of the region (Suter, 2014). Another aspect of localization that deserves special attention is the time frame in which marketing choices are made. Proactive marketing is where the company plans localization before the product’s foreign distribution to achieve an established marketing objective (Suter, 2014). On the other end is retroactive marketing, where marketing plans and localization are developed after distribution, influenced by how the foreign audiences interpreted a certain good (Suter, 2014). A prominent case that balances both of these types of marketing is the Pokémon Company and its anime distribution. When the Pokémon anime was first broadcasted, the proactive marketing choice was made assuming that the Japanese will be its only audience, which may also explain the increased Japanese cultural markers included in those seasons (Raes, 2018). Retroactive marketing appears in its later global distribution leading the company to consider a variety of audiences in both their localization of existing episodes and production of future episodes (Raes, 2018). Retroactive marketing choices echo the anime’s growing international reach, its contents demonstrating a transition from using heavy Japanese cultural imagery in the earlier seasons to diminishing the Japanese cultural odor to include more international influences in later seasons (Raes, 2018). The Company had to balance its initial Japanese audiences, those who held expectations for how anime represents Japanese culture, non-Japanese audiences who are tangentially aware of Japanese culture, and international audiences who are entirely unfamiliar with Japanese culture, (Raes, 2018). To appeal to all of these audiences, the Pokémon Company chose to maintain Pokémon’s Japanese odor through the inclusion of subtle cultural imagery (e.g., featuring Japanese foods) while names are changed to fit local languages and customs, as well as providing multilingual dubbing (Raes, 2018). As such, new audiences unfamiliar with symbols can also comprehend the episodes while the remaining aspects of Japanese odor made the show culturally unique (Raes, 2018).

Effect of Tourism on Local Communities

Tourism uses intentional identity-based marketing to promote products and services. In investigating the effects of this intentional commodification, existing literature depicted a study that was conducted for its effects on both the producers and consumers. Said study presents that providing incentives, in this situation the monetary gains from tourism, undermines an identifier’s intrinsic motivation to distribute their culture and arouses reactance and suspicion within the consumer (Bhattacharjee, 2010). Reactance is defined as the “motivation to restore a threatened freedom” (Bhattacharjee, 2010). In the consumer’s situation, the threatened freedom may be their perceived agency in making purchasing choices, while for the producer of the cultural good and identifier of the commodified culture, it is the freedom to present their culture in a way that aligns with their self-perception. Consumers who identify with the marketed identity may display reactance by de-emphasizing the targeted identity from their self-construct which leads to implications for subsequent, unrelated identity-relevant decisions (Bhattacharjee, 2010).

Beyond the individual level, tourism also greatly impacts lifestyle in both the consumers’ and home culture’s communities. For tourists, traveling incites social change as it invites the consumer to experience novelty and possibly bring back such change to their hometown (Moeran, 1983). Similarly, communities frequented by tourists are evidenced to also adapt their lifestyles to accommodate the influx of foreign visitors. One 2017 study on the effects of tourism made findings by focusing on the case of Niseko, a small skiing town in Japan that has experienced rapid growth in international tourism (Nelson and Matthews, 2017). There was a generally positive response from local residents towards the influx of tourists due to the financial contributions they made to the town, with ambivalent comments towards lifestyle changes they would need to promote tourism such as the increased need to learn English and foreign property ownership (Nelson and Matthews, 2017).

3. Discussion

The literatures that were consulted are put together into a framework anticipating how culture commodification changes self-perception and corresponding behavior. This framework focuses on three different aspects—individualist vs. collectivist cultures, high vs. low identifiers, and the presence of monetary benefit—and how they combine in an interactive effect.

3.1 Individualist vs. Collectivistic Cultures

The following sections discuss how individualist and collectivist cultures alter their self-perception by demonstrating reactance in response to cultural commodification. To reiterate, reactance is defined as “the motivation to restore a threatened freedom,” and in the context of cultural identification and commodification refers to the identifier’s perceived identity threat upon realizing the disparity between their cultural self-perception and how non-identifiers view their culture (Bhattacharjee, 2010).

Based on the research examined, this paper proposes that collectivists will demonstrate reactance towards products that represent cultural commodification if they feel that their community experiences identity threat from that particular product. Said reactance may mainly manifest as a negative attitude towards the product representing that commodification, namely an avoidance to purchase or produce it. For example, if a Japanese company produces kimonos that are rented out to tourists for photo-taking, but the design of those kimonos is historically inaccurate as they have been simplified for an easier wear, Japanese people will avoid renting this brand of kimonos even if the same brand has a separate line that is historically accurate. However, this case of the perceived threat from commodification does not stop them from engaging with other similar services. If a different company produces accurate kimonos for tourists to rent, Japanese people may be more inclined to still engage with that service as they do not find that case of commodification threatening their identity. This selectivity in the tendency to formulate reactance on a product-by-product basis is constructed from previous studies that have indicated collectivists rely on situational cues for judgment (Cousins, 1989).

Additionally, as they are also less egocentric, they will not maintain a self-consistent negative attitude towards all products if there are relevant others who favor some (Cohen and Gunz, 2002). Collectivists will respond with reactance to protect their perception of authenticity within their displays of culture, and if anyone in the collectivist’s relevant group does not sense a threat to authenticity, they will not adopt reactance.

The opposite argument applies to identifiers of individualist cultures. Due to their more egocentric, self-consistent nature, they are likely to generalize their sentiments across all cases of commodification without being influenced by what others in their community think (Cohen and Gunz, 2002; Kim, et al., 2014). While the individualist’s attitude formation may be guided by those around them, due to their inborn self-consistency, once an individualist’s opinion solidifies for commodification it will be relatively difficult to change it (Kim, et al., 2014). Thus, if an individual has a predisposed negative attitude towards commodification, they will display negative attitudes towards all cases of commodification for their culture regardless of the potential threats or benefits it might provide. Similarly, if an individual has a predisposition to favor commodification, they will also display positive attitudes toward all cases of commodification in their culture.

The discussions of individualist and collectivist cultures are summarized into two propositions:

P₁: Collectivists are more likely to experience and demonstrate reactance based on singular commodities’ assessed threat to their community

P₂: Individualists are likely to generalize their sentiments across all cases of commodification, holding a positive or negative attitude towards all goods that commodify their culture rather than holding varying opinions between each.

3.2 High vs. Low Identifiers

This portion aims to hypothesize the mindset and behaviors different identifiers will take on after experiencing the identity threat that comes as a result of commodification. While this is not a distinction clearly established within the previous literature

consulted, this paper proposes to divide identity-threat into two categories within this analysis: devaluation—the threat one feels when their culture is belittled or devalued by commercialization—as well as distinctiveness loss—the threat one feels in the face of assimilation. As respective examples, one may feel threatened if a product labeled as being from their culture is priced less compared to a similar product labeled to be from a different culture; one may also feel threatened if the local language were spoken less due to a large influx in tourism.

When cultural identities are threatened by devaluation during commodification, higher identifiers of a culture are more likely to continue affirming their connection while lower identifiers are likely to detach their connection to the identity post-commodification. This statement is supported by the associative and dissociative behavior linked with independent and interdependent self-construals as well as collective self-esteem, all of which have been mentioned while discussing high vs. low identification in the literature review (White, et al., 2012; White and Argo, 2009). While the study on CSE specifically referred to the brand-identity linkage where association and dissociation with identity are purely demonstrated through the purchasing of identity-related products, this proposal intends to include all forms of identity expression, whether it be cognitive or behavioral. For instance, if a higher identifier of Chinese culture observes that Chinese food is less preferred in their community though they enjoy it themselves, the identifier will not only still be as inclined before the observation to consume Chinese food but will also not feel or express that they are any less Chinese than before. In the same situation, a lower identifier may feel less motivated to express liking or consuming Chinese food after observing the value difference and be less willing to think of themselves as Chinese and express their Chinese identity.

While devaluation may be defended through passive associative or dissociative response, identity threat through distinctiveness loss can evoke aggressive defensive behavior titled closure, the active display of negative attitudes towards foreign influence (Morris, et al., 2011). This paper proposes

that higher identifiers will protect the perception of the culture they hold pre-commodification, and demonstrate closure towards aspects of their culture disrupted by distinctiveness loss post-commodification. To explain with application, if Singapore gains an increasing Chinese population due to an increase in tourism or immigration, what constitutes Singaporean culture will certainly change to include more Chinese influence, similar to a culture post-commodification. However, higher identifiers of the original Singaporean culture will still call themselves Singaporean in alignment with their image of that community being defined by the Singapore pre-Chinese influence. However, they will also display closure by performing traditional Singaporean activities, particularly those that would exclude Chinese immigrants (e.g., speaking in Singlish rather than standard English which Chinese immigrants tend to understand more), thereby demonstrating the higher identifier's possibility of maintaining a static self-perception of culture post-commodification and displaying closure.

The discussions of high vs. low identifiers are summarized into two propositions:

P₃: In response to identity threat through the devaluation of a culture, higher identifiers will continue affirming their cultural connection across all cognitive and behavioral expressions of the culture.

P₄: When threatened by distinctiveness loss through cultural commodification, higher identifiers' cultural perceptions will

a. remain static from their pre-commodification perception.

b. demonstrate closure towards content they perceive as being part of the post-commodification culture.

3.3 Presence of Monetary Benefit

Monetary benefit has been evidenced to undermine intrinsic motivation by overriding it with extrinsic reward (Bhattacharjee, 2010). By removing intrinsic motivation, monetary benefit then elicits reactance within the consumer as it removes their freedom of choosing something for themselves without being swayed by external factors (Bhattacharjee, 2010). In the context of identity-based marketing, a consumer will then

demonstrate reactance by de-emphasizing the identity from their self-perception (Bhattacharjee, 2010). By applying this detachment to the producer's side, one can propose that the reactance elicited by monetary benefit would then cause identifiers to separate the idea of their own culture from the idea of the aspect of their culture that they are outputting or commodifying to a foreign audience. The construction of two separate cultural perceptions targeting the in-group and out-group by an identifier may then explain situations such as the difference between what is provided to Japanese fans and foreign fans of anime. For example, the level of cultural odor produced in anime that are catered to Japanese audiences is much higher than those catered towards foreign audiences, as in the example of earlier vs. later seasons of the Pokémon anime (Raes, 2018). The later seasons of Pokémon lower in cultural odor reflect the separate brand and cultural identity the Japanese company developed to cater to a foreign audience and out-group, by extension indicating the marketing choice as a behavioral manifestation of the changed self-perception commodification has caused. Another disparity between the displays of Japanese culture towards the domestic vs. international audience is in the lack of offerings for anime-related consumption. Though the market for anime merchandise is very well-developed in Japan with entire districts such as Akihabara being dedicated to their selling, research demonstrates that anime is rarely advertised in promotional materials for Japan aimed at foreign travelers (Sabre, 2017). These two examples respectively reflect that Japanese cultural odor in anime and media-based consumption, while being included in the domestic perception and presentation of Japanese culture, is not found in the Japanese people's views of what Japanese culture towards an out-group should look like.

Not only is the aspect of a culture expressed and perceived by an identifier different from the one of a non-identifier, but the presence of monetary benefit post-commodification also causes an identifier to describe the outputted aspects of their culture with more stereotyping than they would for the internal aspects of their culture. The definition of stereotypes here is not necessarily stereotypes that are specific to a culture, but rather the identifier's stereotypes of

what aspects of their culture the foreign audience would want to engage with, and how they would want to engage with it. Hence, the Japanese, despite being a collectivist culture, will choose to emphasize individuality and *kosei* in their travel brochures under the assumption that they are appealing to tourists from individualist cultures, as well as learning English to improve communication with tourists as is the case of Niseko (Moeran, 1983; Nelson and Matthews, 2017). These stereotypes then result in the disparity between tourist demands and supply, creating situations such as in Sabre's 2017 study where French tourists to Japan demonstrated interest in anime and Japanese popular media, but failed to find such content advertised in tourist brochures (Sabre, 2017).

The discussions of monetary benefit can be summarized into two propositions:

P₅: Reactance elicited by monetary benefit causes individuals to separate the idea of their own culture from the aspects of their culture that they are exporting to a foreign audience.

P₆: Monetary benefit causes individuals to describe their culture's interactions with another culture with more stereotyping than they would for describing their in-group perception.

3.4 Interactive Effects

While the previous proposals examine each factor's individual effect on cultural perception post-commodification, this portion aims to outline claims that discuss the components' effects alongside each other.

Firstly, higher identifiers of individualist cultures are predicted to have an unchanged perception of their culture after commodification. The static nature of their cultural self-perception is credited to the individualist's egocentric and self-consistent nature (Cohen and Gunz, 2002; Kim, et al., 2014). If a high identifier of individualism already has a preconceived notion of what their culture should be, those two traits make it more likely for them to maintain those beliefs and therefore unlikely to change their cultural perceptions post-commodification. Additionally, even if they do eventually align their perceptions with the post-commodification perception they are likely to

view the change as one of their own choice rather than as a product of external force due to the individualist's tendency to rationalize away conformist behavior (Hornsey and Jetten, 2005).

High identifiers of individualism will also display closure through the rejection of all commodification if the perception of their culture is challenged, even if it is only by one product of the aftermath of commodification. This generalized reactance is explained by individualists' overall tendency to make larger inferences from singular events (Cousins, 1989).

Similar to the high identifiers of individualist cultures, low identifiers will also maintain an unchanged perception of their culture and resist challenges to that image due to their core individualism. However, the presence of monetary benefit is what sets the two apart. Individualists are likely to continue producing goods that commodify their culture as they will prioritize the monetary benefits for the self over the spread of inaccurate narratives the collective will experience. However, as stated before, the monetary benefit commodification has provided arouses reactance (Bhattacharjee, 2010). For the higher identifier, such challenges are dealt with through associative means while for the lower identifier dissociative means (White and Argo, 2009; White, et al., 2012). Applying the differing ways identifiers display reactance to the continuation of production individualists will have to gain monetary benefit for the self, high identifiers of individualism may utilize egocentric projection and association by producing more products that align with their cultural self-perception, while the lower identifiers of individualism will dissociate by continuing to produce cultural commodities but not incorporating them as part of their cultural self-perception. To illustrate the latter part with an example, if commodification made foreign audiences associate the bald eagle with the United States but a low identifier of the United States does not agree with that, they will continue producing hats labeled as being from the United States that contain imagery of the bald eagle but do not view the hats as being a cultural product themselves.

Similar to the low identifiers of individualist cultures, higher identifiers of collectivist cultures will

also maintain the same perception of their culture post-commodification. However, they will continue producing cultural products that are in alignment with the post-commodification image though they have a pre-commodification perception. As members of collectivist culture tend to have more context-dependent perceptions of self, they are likely to capitalize on that in their reactance to monetary benefit by developing the aforementioned two different views of their culture - one aspect geared towards the in-group and one towards the out-group (Cousins, 1989; Bhattacharjee, 2010). As such, the aspect of their culture they interpret to be for the in-group will be what remains unchanged after commodification, while the self-perception of their culture they developed for the out-group will follow the changes commodification ensued and thus lead to increased production of cultural goods that reflect the change done by commodification.

Lower identifiers of collectivism, being somewhat fluid in their self-perceptions due to their collectivist core but not having a strong definition of their culture due to their low identification, are likely to alter their perception of their culture according to the changes of commodification. However, as the production of cultural goods will still elicit reactance, it will cause the lower identifiers to exhibit a dissociative response which is demonstrated through the production of culturally odorless products.

The interactive effects are summarized into four propositions:

P₇: Higher identifiers of individualist cultures will display closure through the aggressive rejection of commodification if their perception is challenged.

P₈: Lower identifiers of individualist cultures will display passive, avoidant behavior to cultural commodities that challenge their perception.

P₉: Only lower identifiers of collectivist cultures will alter their perception according to the changes commodification causes but choose to produce more culturally odorless products instead.

P₁₀: Higher identifiers of collectivist culture will produce more products that cater to the altered image for monetary benefit.

4. Conclusion

In summary, the paper has reviewed and listed the various cognitive and behavioral responses people may have to a discrepancy between their own and others' views of their cultures post-commodification. After examining subsets of relevant literature, this paper is able to construct propositions based on the independent effects of cultural individualism, level of identification, and salience of financial gain, as well as the interactive effects between the first two factors on the latter that altogether form a framework for predicting an individual's reactions to cultural commodification.

While the literature examined here is only a subsection of all works and can nowhere represent the full variety of conclusions made across the entire field, all proposals have been made with analysis and cross-referencing between experimental studies and reviews, as well as its applicability to actual case studies.

4.1 Contributions

This investigation exists as the pioneer of an intersectional field that combines the intrinsic characteristics of certain cultures and identifiers as well as extrinsic motivation and applies their grounds to multiculturalism rather than investigating a singular community. The framework developed is unique from two perspectives. Firstly, the proposals are centered around an individual's reactions to cultural commodification rather than a group or a culture, creating a connection between macro phenomena such as changes in a product's marketing with micro aspects such as one's change in perception. Secondly, the framework explores cultural marketing with respect to the identifiers or producers of the commodified culture rather than the foreign consumers. It is hoped that the framework provided in this review will direct cultural businesses towards authentic representation, without the need to compromise their originality for monetary gain.

4.2 Areas for Future Research

It is of practical and theoretical significance to discuss whether this work is applicable to advance research in cultural interaction and self-perception,

particularly regarding commodification and globalization. Future research may work to verify the framework proposed. New studies can be conducted regarding the relationship between cultural odor and reactance, more concretely between cultural products and cultural schema as to whether identifiers do consider certain cultural products to be part of their culture. Additionally, one can investigate whether certain cultural media or products elicit more reactance or change odor more during commodification to foreign audiences. Returning to the exploration of self-perception, the process by which cultural products are developed in relation to cultural self-image can also be researched. Lastly, expanding on monetary benefit, future research may also address whether there exist other forms of extrinsic motivation that affect commodification.

References

- Amarnath, D. D., & Jaidev, U. P. (2020). Toward an integrated model of consumer reactance: A literature analysis. *Management Review Quarterly*, 71(1), 41-90. <https://doi.org/10.1007/s11301-020-00180-y>
- Bauer, H. H., Neumann, M. M., & Huber, F. (n.d.). Intercultural consumer classifications in e-commerce. *Studies in Classification, Data Analysis, and Knowledge Organization*, 561-568. https://doi.org/10.1007/3-540-28084-7_66
- Bhattacharjee, A. (2010). Constraints and Consequences: Psychological Reactance in Consumption Contexts. *Association for Consumer Research*, 37, 123-126.
- Briley, D. A., Morris, M. W., & Simonson, I. (2022). Reasons as Carriers of Culture: Dynamic versus Dispositional Models of Cultural Influence on Decision Making [PDF]. *Journal of Consumer Research*, 49(1). https://dpl6hyzg28thp.cloudfront.net/media/Reasons_as_Carriers_of_Culture.pdf
- Clee, M. A., & Wicklund, R. A. (1980). Consumer Behavior and Psychological Reactance. *Journal of Consumer Research*, 6, 389-405.
- Cohen, D., & Gunz, A. (2002). As seen by the other

... : perspectives on the self in the memories and emotional perceptions of easterners and westerners. *Psychological Science*, 13(1), 55-59. <https://doi.org/10.1111/1467-9280.00409>

Cool Japan Proposal. (n.d.). Japanese Cabinet Office. Retrieved August 2, 2022, from https://www.cao.go.jp/cool_japan/english/pdf/published_document3.pdf

Cousins, S. D. (1989). Culture and Self-Perception in Japan and the United States. *Journal of Personality and Social Psychology*, 56, 124-131.

Giacomin, M., & Jordan, C. (2017). Interdependent and independent self-construal. *Encyclopedia of Personality and Individual Differences*, 1-7. https://doi.org/10.1007/978-3-319-28099-8_1136-1

Giunta, C. (2020). Digital Marketing Platform Tools, Generation Z, and Cultural Considerations [PDF]. *Journal of Marketing Development and Competitiveness*, 14(2), 63-75. http://digitalcommons.www.na-businesspress.com/JMDC/JMDC14-2/6_Guinta_14_2_.pdf

Hammou, I., Aboudou, S., & Makloul, Y. (2020).

Social media and intangible cultural heritage for digital marketing communication: Case of marrakech crafts. *Marketing and Management of Innovations*, (1), 121-127. <http://doi.org/10.21272/mmi.2020.1-09>

Hausmann, A., & Weuster, L. (2017). Possible marketing tools for heritage tourism: The potential of implementing information and communication technology. *Journal of Heritage Tourism*, 13(3), 273-284. <http://dx.doi.org/10.1080/1743873X.2017.1334786>

Hirshberg, M. S. (1993). The Self-Perpetuating National Self-Image: Cognitive Biases in Perceptions of International Interventions. *Political Psychology*, 14(1), 77-98. <https://doi.org/10.2307/3791394>

Holland, J., & Gentry, J. W. (1999). Ethnic consumer reaction to targeted marketing: A theory of intercultural accommodation. *Journal of Advertising*, 28(1), 65-77. <https://doi.org/10.1080/00913367.1999.10673577>

Hornsey, M., & Jetten, J. (2005). Loyalty without conformity: Tailoring self-perception as a means of balancing belonging and differentiation. *Self and Identity*, 4(1), 81-95. <https://doi.org/10.1080/13576500444000236>

Novel Therapies in the Treatment of Parkinson's Disease

Eshaan Gandhi^{1*}

¹Indian Hill High School, Cincinnati, OH USA

Received October 17, 2022; Revised November 18, 2022; Accepted, November 28, 2022

Abstract

Parkinson's Disease (PD) is the second most common neurodegenerative disease, caused by the degenerative progression of dopaminergic neurons located in the substantia nigra pars compacta, which results in the decreased production of dopamine — a necessary neurotransmitter to maintain homeostasis. However, the ultimate cause of damage done to these neurotransmitters is unknown, as it has been found that the development of Lewy Bodies, overproduction of calcium ions, and other protein mutations can all cause dopaminergic neurons to be deficient. The development of PD causes deteriorating symptoms such as bradykinesia, motor tremors, muscle stiffness, and impaired balance. No cure for Parkinson's disease has been developed, in which researchers have instead developed therapies to combat the degenerative symptoms that result from PD. Recently, there has been an amplification in research and generation of medical pharmacologic therapies and invasive surgical therapies, such as Deep Brain Stimulation, to alleviate symptoms. Although current pharmacologic therapies function to a certain degree, they are not effective over long periods, which is how PD becomes increasingly damaging. Similarly, there is a lack of definitive disease-modifying therapies such as alpha-synuclein antibodies, LRRK2 antibodies, exenatide, and isradipine. These disease-modifying pharmaceuticals, surgical therapies, and future treatments must continue to be researched for safety, longevity, and effectiveness to promote optimal quality of PD patients.

Keywords: Parkinson's Disease, Therapy, Agents, Mutation, Neuroscience, Neurodegenerative Disease

1. Introduction

Parkinson's Disease (PD) is a progressive neurodegenerative disorder that is caused by the loss of specific neurons within the substantia nigra pars compacta, which are responsible for the production of dopamine. Dopamine is extremely important to the well-being of an individual, as it serves as a chemical messenger between various parts of the brain and nervous system to coordinate bodily movement (Kaila & Lang, 2015). As dopamine production decreases, the ability of an individual to control and coordinate bodily movements with ease progressively deteriorates. PD has been around for a long period but was first discovered by an English physician

named James Parkinson, who the disease is named after. Today, PD is considered the second most common neurodegenerative disease after Alzheimer's Disease, affecting more than 10 million people worldwide. Although PD affects all cultures and races equally, it is approximately 2 times more prevalent in males than in females (Cerri et al. 2019).

James Parkinson first described the disease as “shaking palsy” after observing the tremors that could occur in someone with the disease. In addition to these tremors, people with PD may encounter coordination trouble, slow movement, deteriorating posture and balance, and stiffness in their legs, arms, or neck. Because of these degenerative symptoms that are almost guaranteed to develop in an individual

* Corresponding Author
eshaan.gandhi24@ihdsd.us

Advisor: Sarah Croston
sarah.croston@ihdsd.us

with PD, many people want to treat the disease. Unfortunately, not only are there no specific prevention techniques for PD, but there are no definitive treatment plans that can eradicate the neurodegenerative disease from the body. Instead, various types of intensive therapies have been developed to combat the grueling symptoms that come with PD that normally result from a mutation in specific genes, ultimately making certain proteins dysfunctional and harmful towards dopaminergic neurotransmitters. Currently, the therapies that exist are somewhat effective, but have many problems that leave a lot to be wanted by researchers, neurologists, and patients. In this literature review, we will discuss the effectiveness and problems with current treatments, as well as the effectiveness, extended benefits, and problems with upcoming disease-modifying therapies, mainly including surgical and neural pathway targeting agents. Although PD might be a disease with no cure at present, it is important for researchers exploring these current therapies of PD to understand the problems with these therapies, as well as what upcoming benefits and problems disease-modifying therapies can offer in order for further advances in PD therapies to be developed.

2. Materials and Methods

This paper was created through a literature review, an exploration of academic papers and clinical trials. We methodically searched PubMed and Google Scholar to identify literature reviews, clinical trials, and randomized control trials examining baseline information on different protein mutations that cause PD, as well as associations between disease-modifying therapies and PD. No date or language restrictions were used, although most papers that were used are from the 21st century. The search strategy for these papers included keywords “Parkinson’s” AND “Disease modifying therapies (Alpha-synuclein, LRRK2, Levodopa, COMT-Inhibitor, MAO Inhibitor, Exenatide, Insulin, Isradipine, L-type calcium, or DBS)” OR “Epidemiology”. When more than one clinical trial on the same therapeutic was eligible, the clinical trial with the most participants was used for statistically

significant reasoning.

3. Results and Discussion

3.1 Current Therapies

Levodopa

Levodopa is a central nervous system agent that is effective as a dopamine-replacement therapy. The main use of Levodopa is to relieve bradykinetic symptoms of idiopathic PD as well as be utilized as a baseline drug to which other potential drug therapies for PD are compared. With oral inhalation and infusion via nasogastric tube being the exclusive methods of intake, Levodopa can enter the body and bypass the blood-brain-barrier, unlike neural-produced dopamine, and is converted into dopamine by the aromatic amino-acid decarboxylase enzyme in the central and peripheral nervous system (Figure 1) (Fahn et al., 2004). In a clinical trial determined on evaluating the effectiveness of Levodopa against worsening symptoms in patients with PD, it was found that the progression of PD in patients who received Levodopa treatment rather than the placebo treatment had significantly slowed based on the Unified Parkinsons’ Disease Rating Scale (UPDRS) (Fahn et al., 2004). Although this study constructively evaluated the effectiveness of Levodopa, it failed to evaluate the effectiveness of the drug against significantly long periods. To address this limitation, a prospective study over a mean treatment duration of 11.1 years focused on evaluating the effectiveness of Levodopa on motor control in patients with idiopathic PD determined that motor response improved over the course of the first 5 years of treatment. Subsequently, the motor response of the patients fluctuated in parallel with the conservation of motor response (Alty et al., 2019). This means that compared to the immediate improvement of symptoms caused by the intake of Levodopa, the drug is not as effective in the long run. This may occur because of the progressively deteriorating neurological functionality that PD causes, ultimately leading to gradually worsening symptoms over time. However, based on our observations and examination, we suggest that currently, Levodopa is likely the most reliable and effective treatment for Parkinson’s Disease especially

with the assistance of enzymatic materials. In the future, however, technological advancements may cause Levodopa to become less frequently used.

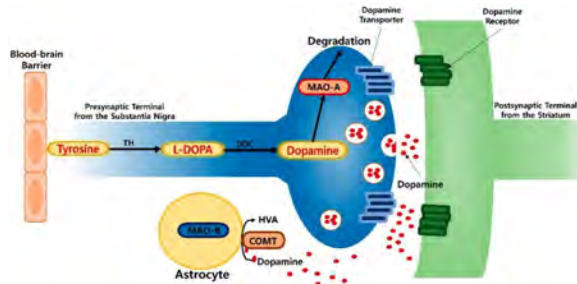


Figure 1. Levodopa's function within individual neurons. Levodopa bypasses the blood-brain barrier and interacts with the synaptic cleft with the assistance of enzymes such as COMT and MAO-B to be converted to dopamine (Jung et al. 2019).

COMT-Inhibitors

Catechol-O-methyltransferase (COMT) inhibitors are a class of drugs used alongside Levodopa to alleviate the worsening motor symptoms associated with PD. Although Levodopa is considered the most effective medication therapy for PD, it loses its effectiveness over time. COMT-Inhibitors can be useful in this sense as they are not only able to extend the effectiveness of Levodopa when used, but they can also allow for lower doses of Levodopa which can have medical and financial benefits for patients (Connolly & Lang, 2014). COMT-Inhibitors function by blocking the further activity of COMT, which is an enzyme that destroys neurotransmitters such as dopamine, the main cause of PD (Finberg, 2019). The most common COMT-Inhibitors used for PD therapy are entacapone, tolcapone, and opicapone, which all slow down the breaking down of Levodopa, allowing more of it to bypass the blood-brain-barrier and reach the brain, leading to the alleviation of PD symptoms. A clinical trial attempted to test how effective COMT-Inhibitors are in combination with Levodopa for PD therapy and found that the use of entacapone — a type of COMT-Inhibitor — was significantly more effective than using Levodopa exclusively in treating motor malfunctions (Merello et al., 1994). The problem with COMT-Inhibitors, however, is that they are not effective as a therapeutic when acting alone (PD Med Collaborative Group et al., 2022). This may ultimately cause patients to have to

purchase more medication and experience more side effects that come with consuming them.

MAO-Inhibitors

Similar to COMT-Inhibitors, Monoamine Oxidase B (MAO-B) Inhibitors are a class of drugs to help treat the symptoms of PD that are usually used alongside Levodopa. However, MAO-B Inhibitors may be used exclusively in the early stages of PD, in which they have moderate effectiveness against PD symptoms. Monoamine Oxidase is an enzyme found in cells and can be classified as either type A or type B. MAO-B is typically found in the brain in which it plays a role in the breakdown of neurotransmitters such as dopamine, ultimately causing the loss of dopamine activity and the appearance of PD symptoms. The goal of MAO-B Inhibitors is to block the activity of this enzyme so these neurotransmitters are lost to a lesser extent (Finberg, 2019). MAO-B Inhibitors have also been found to slow the breakdown of Levodopa after it passes through the blood-brain barrier, which may boost the effectiveness and sustainability of the medication (Finberg, 2019). Typically, MAO-B Inhibitors come in the form of orally consumed tablets, making the effects of the drug easy to obtain. A clinical trial was conducted to test the effectiveness of MAO-B Inhibitors in patients with PD, in which it was found that there were very small yet persistent alleviating effects found in patient mobility scores when MAO-B Inhibitors were used in combination with Levodopa, which is the most common method of use (PD Med Collaborative Group et al., 2022). The study, however, also found that MAO-B inhibitors were not effective in alleviating PD symptoms when used alone, which is the main issue in the utilization of the drug (PD Med Collaborative Group et al., 2022). Ultimately, this causes financial and medical complications, which is why MAO-B Inhibitors aren't the most common nor effective medication in PD symptom relief.

3.2 Molecular Pathway Targeting Agents

Alpha-Synuclein

Alpha-Synuclein is a neural transmitter whose normal function has not been established, but has been suggested to be a cause of familial PD. Mutated alpha-synuclein proteins are thought to cause PD by abnormally aggregating into the form of Lewy Bodies and Lewy Neurites. As a result, the

accumulated alpha-synuclein proteins can spread by binding their fibrils to neuronal cell surfaces to gain entry into the cytoplasm until the entire cell surface of neurons is bound by alpha-synuclein fibrils in which the proteins can overtake the cells (Hijaz et al., 2020). As the proteins spread among neuronal cell pathways, they will eventually overexpress themselves in which they will be able to inhibit and slow neurotransmitter releases, specifically dopamine ultimately leading to neurodegeneration. To combat the spread of alpha-synuclein in cellular pathways and host cells, advances in therapies have been developed to produce disease-modifying effects by targeting the spread, production, aggregation, and degradation of alpha-synuclein. We suggest that it is important for scientists to examine therapy alpha-synuclein further, as it is a mutation that a significant amount of people have, risking these individuals for PD. An example of these therapies can be found in a clinical trial investigating the tolerability and safety of alpha-synuclein antibody PRX002 over 2 years, where it was determined that repeated treatment of PRX002 was generally safe and tolerable for patients with PD while inducing reductions in free serum alpha-synuclein (target engagement) (Jankovic et al., 2018). Unfortunately, little is known about the holistic effectiveness of these antibodies and therapies against the aggregation of alpha-synuclein in molecular pathways, so more experimental research is needed to determine whether they are realistic approaches to combat PD symptoms.

Leucine-Rich Repeat Kinase 2

Leucine-Rich Repeat Kinase 2 (LRRK2) is a large protein that has many different functions but has also been suggested to account for 5-13% of familial PD and 1-5% of sporadic PD when mutated (Qin et al., 2018). When the LRRK2 protein mutates, it causes the protein's enzymatic activity to uncontrollably increase kinase activity in the kinase domain or through the GTPase domain by intentionally weakening and damaging the GTP hydrolysis process (Nguyen et al., 2020). With this, neuronal toxicity is induced in cultured cells which has been suggested to impair dopaminergic activity, ultimately leading to the development of PD over time. It has also been

determined that mutated LRRK2 may influence the aggregation of alpha-synuclein, ultimately invading cell structures and molecular pathways through increased kinase activity (Zimprich et al., 2004). To combat these imminent mutative threats, certain therapeutic interventions have been developed to prevent neurodegenerative effects. One of these developments are LRRK2 inhibitors, which were tested in preclinical trials and were suggested to prevent some aspects of neurodegeneration by potentially using autophagy to replace deteriorating cells (Lee et al., 2012). Although this trial was mostly successful, it was limited by lung toxicity which could cause other harms when given, which is why more research needs to be conducted on this treatment type. In fact, treatment of LRRK2 mutation is very much incomplete and dangerous as it can have detrimental effects on its victims.

3.3 Deep Brain Stimulation

Deep Brain Stimulation (DBS) is a minimally invasive procedure that has the potential for symptom alleviation in those with neurodegenerative diseases, most commonly, Parkinson's Disease. To start DBS, a thin wire covered in electrodes is implanted to induce high-frequency electrical stimulation in specific regions of the brain. The location where the electrodes are placed depends on what area of the brain is creating abnormal signals which are causing the patient to experience symptoms. To power the signal-inducing electrodes, a battery-powered operator is connected to the thin wire, and placed into the chest wall. The battery-powered operator is regulated by a computer, which contains instructions for when the electrodes should send electrical impulses. When properly functioning, the electrodes send pulses of electrical signals to normalize the occurring abnormal electrical activity in which the defective signals coming from likely damaged neurons are blocked (Fariba & Gupta, 2021). To test the effectiveness of DBS in alleviating the motor deficiency symptoms of PD, a clinical trial was conducted where it was found that neurostimulation of the subthalamic nucleus was significantly more effective than solely medication where UPDRS scores had mean improvements of 9.5-19.6 points

(Deuschl et al., 2006). Although DBS offers some beneficial effects for those with neurodegenerative symptoms, there is a significant risk for extremely damaging adverse events. Examples of such adverse events include fatal internal brain hemorrhages, device malfunctions, marginal paralysis, reduced coordination, and deteriorating emotional and mental status (Topp et al., 2021). With the potential for patients to undergo these adverse events, more research and testing must be conducted to reduce the limitations associated with DBS. This is mainly because of the beneficial effects that DBS offers when successful in which it can alleviate PD symptoms more effectively than most medications. After examining the effectiveness of DBS, we may suggest that in present day, DBS is not the safest nor reliable treatment for Parkinson's Disease. It is likely that this may not be the case in the future due to technological advancements in this field.

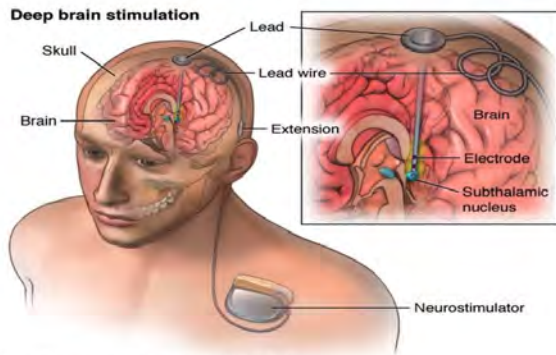


Figure 1 Anatomical illustration of DBS STN

Figure 2. Diagram showing how Deep Brain Stimulation functions as a treatment, as well as materials used in their specific locations. Electrodes placed in the brain are connected to a neurostimulator, which sends electrical signals to stimulate electrical activity within the brain (Mayo Foundation, 2021).

3.4 Neuronal Rescue Pathway Targeting Agents

Exenatide

Glucagon-like peptide-1 (GLP-1) receptor agonists are medications that mimic a naturally occurring hormone that can have neuroprotective effects against PD while typically being used as a medication to treat Type-2 diabetes. The most tested form of these agents is Exenatide, which is typically used to help an

individual's pancreas produce insulin more efficiently to control blood sugar levels better. However, in a recent clinical trial investigating the effectiveness of Exenatide against motor PD, it was determined that Exenatide had beneficial effects on motor scores that sustained beyond the period of exposure to the medication (Athauda et al., 2017). Not only this, but in an open-label Phase II study, it was determined that not only did deteriorating symptomatic ataxia improve, but cognitive-behavior improved as well (Aviles-Olmos et al., 2013), making Exenatide an extremely beneficial impermanent medication for PD. Exenatide and other GLP-1 receptor agonists are so effective because they have been shown to be beneficial to mechanisms regulating reduced inflammation that appears after the development of PD as well as combative to continued aggregation of alpha-synuclein as seen in animal models (Wang et al., 2021). Although these medications are effective, more research must be conducted to determine if Exenatide can be used as a long-term medication for deteriorating developments that occur during progressional PD.

L-Type Calcium Channel Blockers/Isradipine

Recent studies have suggested that dihydropyridine calcium-channel blockers, which are typically used to treat hypertension and angina, may be associated with a reduced risk of developing deteriorating symptoms in PD. This is because certain neuronal areas with spontaneous peacemaking cells, such as substantia nigra pars compacta, rely on L-type calcium channels that allow increased calcium entry into cells. With increased calcium entry, the body is more at risk for increased oxidative stress, mitochondrial damage, and increased cell death (Chan et al., 2007). To continuously combat this issue, Isradipine is typically used as a calcium-blocking medication. To test the effectiveness of Isradipine, a clinical study was conducted in which results determined that short-term treatment of PD using Isradipine may be effective, but long-term usage does not slow the progression of PD which acts as a limiting factor for this medication (Parkinson Study Group STEADY-PD III Investigators, 2020). Contrastingly, in a statistical analysis assessing the effectiveness of non-specific L-type calcium channel blockers on PD, it was found that these medications had a 27% risk reduction in developing ataxic and bradykinetic symptoms (Ritz

et al., 2010). This may suggest that Isradipine, the most common form of L-type calcium channel blocker, may not be as effective as other L-type calcium channel blocker medications.

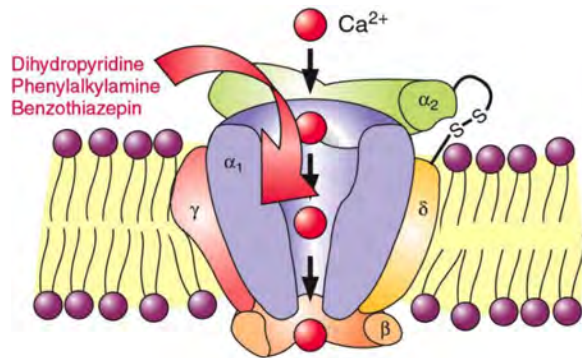


Figure 3. Diagram showing the anatomy of a L-Type Calcium Channel, as well as dihydropyridine calcium-channel blockers entering the channel as inhibitors (Wikipedia 2022).

4. Conclusion

It appears that PD is associated with a plethora of pathophysiological causes, ranging from the aggregation of alpha-synuclein, mitochondrial dysfunction, increased calcium production, and neuroinflammation. All of these causes, however, cause the impairment of dopaminergic activity, which is something that most therapies for PD attempt to fix. Although we haven't developed strategies to completely restore the neurotransmitters producing a lack of dopamine, we have developed therapies that can at least alleviate the symptoms and suffering that patients with PD undergo. To improve these therapeutics and use them to their full potential, we suggest that it will be necessary to advance methods that are used to quantify the extent to which a patient with PD responds to a certain disease-modifying approach, whether that's through molecular pathway targeting agents, levodopa, surgical methods, or neuronal rescue pathway targeting agents. By being able to measure the extent to which these therapeutics are effective, researchers and neurologists may be able to make a better prediction of how people will respond to therapeutics through the alleviation or pursuit of deteriorating motor and gait symptoms. With advances in technology and research, it will be

necessary to not only consider the therapies mentioned in this review, but also to consider the restoration of neurotransmitters using microscopic technology as a potential cure for PD. Finally, based on our examinations while researching this topic, we have come to the conclusion that it would be the effective for people suffering with PD to continue to take Levodopa in combination with MAO-B or COMT Inhibitors. This combination of therapies is not only the most effective treatment for PD in the present day, but it is also the most cost-effective. Overall, although current and upcoming therapies for PD are somewhat effective, we heavily suggest it will be necessary for more developmental changes to occur in the attempt to find a terminal cure for PD.

Acknowledgment

I would like to thank the UCI X GATI program for the opportunity to write this paper, as well as Mrs. Sarah Croston for sponsoring our work. I would also like to thank my family for their continued support in my scientific research journey.

References

- Alty, J., et al. (2009). Longitudinal study of the levodopa motor response in Parkinson's disease: relationship between cognitive decline and motor function. *Movement disorders : official journal of the Movement Disorder Society*, 24(16), 2337–2343. <https://doi.org/10.1002/mds.22800>
- Athauda D., et al. (2017). Exenatide once weekly versus placebo in Parkinson's disease: a randomised, double-blind, placebo-controlled trial. *Lancet* (London, England), 390(10103), 1664–1675. [https://doi.org/10.1016/S0140-6736\(17\)31585-4](https://doi.org/10.1016/S0140-6736(17)31585-4).
- Aviles-Olmos I., et al. (2013). Exenatide and the treatment of patients with Parkinson's disease. *The Journal of clinical investigation*, 123(6), 2730–2736. <https://doi.org/10.1172/JCI68295>
- Cerri, S., Mus, L., & Blandini, F. (2019). Parkinson's Disease in Women and Men: What's the Difference?. *Journal of Parkinson's disease*, 9(3), 501–515. <https://doi.org/10.3233/JPD-191683>
- Chan, C., et al. (2007). 'Rejuvenation' protects

- neurons in mouse models of Parkinson's disease. *Nature*, 447(7148), 1081–1086. <https://doi.org/10.1038/nature05865>
- Connolly, B. S., & Lang, A. E. (2014). Pharmacological treatment of Parkinson disease: a review. *JAMA*, 311(16), 1670–1683. <https://doi.org/10.1001/jama.2014.3654>
- Deuschl, G., et al. (2006). A randomized trial of deep-brain stimulation for Parkinson's disease. *The New England journal of medicine*, 355(9), 896–908. <https://doi.org/10.1056/NEJMoa06028>.
- Fahn, S., et al. (2004). Levodopa and the progression of Parkinson's disease. *The New England journal of medicine*, 351(24), 2498–2508. <https://doi.org/10.1056/NEJMoa033447>
- Fariba, K. A., & Gupta, V. (2021). Deep Brain Stimulation. In StatPearls. StatPearls Publishing.
- Finberg J. (2019). Inhibitors of MAO-B and COMT: their effects on brain dopamine levels and uses in Parkinson's disease. *Journal of neural transmission (Vienna, Austria : 1996)*, 126(4), 433–448. <https://doi.org/10.1007/s00702-018-1952-7>
- Hijaz, B. A., & Volpicelli-Daley, L. A. (2020). Initiation and propagation of α -synuclein aggregation in the nervous system. *Molecular neurodegeneration*, 15(1), 19. <https://doi.org/10.1186/s13024-020-00368-6>.
- Jankovic, J., et al. (2018). Safety and Tolerability of Multiple Ascending Doses of PRX002/RG7935, an Anti- α -Synuclein Monoclonal Antibody, in Patients With Parkinson Disease: A Randomized Clinical Trial. *JAMA neurology*, 75(10), 1206–1214. <https://doi.org/10.1001/jamaneurol.2018.1487>.
- Jung, H. A et al. (2019). Novel Diels-Alder Type Adducts from *Morus alba* Root Bark Targeting Human Monoamine Oxidase and Dopaminergic Receptors for the Management of Neurodegenerative Diseases. *International journal of molecular sciences*, 20(24), 6232. <https://doi.org/10.3390/ijms20246232>
- Kalia, L. V., & Lang, A. E. (2015). Parkinson's disease. *Lancet (London, England)*, 386(9996), 896–912. [https://doi.org/10.1016/S0140-6736\(14\)61393-3](https://doi.org/10.1016/S0140-6736(14)61393-3)
- Lee, B. D., Dawson, V. L., & Dawson, T. M. (2012). Leucine-rich repeat kinase 2 (LRRK2) as a potential therapeutic target in Parkinson's disease. *Trends in pharmacological sciences*, 33(7), 365–373. <https://doi.org/10.1016/j.tips.2012.04.001>.
- Mayo Foundation for Medical Education and Research. (2021). *Deep Brain stimulation*. Mayo Clinic. <https://www.mayoclinic.org/tests-procedures/deep-brain-stimulation/about/pac-20384562>
- Merello, M., et al. (1994). Effect of entacapone, a peripherally acting catechol-O-methyltransferase inhibitor, on the motor response to acute treatment with levodopa in patients with Parkinson's disease. *Journal of neurology, neurosurgery, and psychiatry*, 57(2), 186–189. <https://doi.org/10.1136/jnnp.57.2.186>
- Nguyen, A., et al. (2020). Dopaminergic neurodegeneration induced by Parkinson's disease-linked G2019S LRRK2 is dependent on kinase and GTPase activity. *Proceedings of the National Academy of Sciences of the United States of America*, 117(29), 17296–17307. <https://doi.org/10.1073/pnas.1922184117>.
- Parkinson Study Group STEADY-PD III Investigators (2020). Isradipine Versus Placebo in Early Parkinson Disease: A Randomized Trial. *Annals of internal medicine*, 172(9), 591–598. <https://doi.org/10.7326/M19-2534>.
- PD MED Collaborative Group, et al. (2022). Long-term Effectiveness of Adjuvant Treatment With Catechol-O-Methyltransferase or Monoamine Oxidase B Inhibitors Compared With Dopamine Agonists Among Patients With Parkinson Disease Uncontrolled by Levodopa Therapy: The PD MED Randomized Clinical Trial. *JAMA neurology*, 79(2), 131–140. <https://doi.org/10.1001/jamaneurol.2021.4736>.
- Ritz, B., et al. (2010). L-type calcium channel blockers and Parkinson disease in Denmark. *Annals of neurology*, 67(5), 600–606. <https://doi.org/10.1002/ana.21937>
- Rui, Q., et al. (2018). The Role of LRRK2 in Neurodegeneration of Parkinson Disease. *Current neuropharmacology*, 16(9), 1348–1357. <https://doi.org/10.2174/1570159X16666180222165418>.
- Topp, G., et al. (2021). Safety of Deep Brain

Stimulation Lead Placement on Patients Requiring Anticlotting Therapies. *World neurosurgery*, 145, e320–e325. <https://doi.org/10.1016/j.wneu.2020.10.047>

Wang, V., et al. (2021). Sustained Release GLP-1 Agonist PT320 Delays Disease Progression in a Mouse Model of Parkinson's Disease. *ACS pharmacology & translational science*, 4(2), 858–869. <https://doi.org/10.1021/acspsci.1c00013>.

Wikipedia. (2022). *L-type Calcium Channel*. https://en.wikipedia.org/wiki/L-type_calcium_channel

Zimprich, A., et al. (2004). Mutations in LRRK2 cause autosomal-dominant parkinsonism with pleomorphic pathology. *Neuron*, 44(4), 601–607. <https://doi.org/10.1016/j.neuron.2004.11.005>

Vagus Nerve Stimulation and Right Median Nerve Stimulation in the Treatment of Coma: A Review of Previous Studies, Mechanisms of Action, and Future Potentials

Chloe Y. Chan¹*

¹Lowell High School, San Francisco, CA, USA

Received May 1, 2022; Revised December 12, 2022; Accepted, January 2, 2023

Abstract

Disorders of consciousness affect a large percentage of patients with brain trauma, and diagnosed patients face a high mortality rate. Recently though, two types of nerve stimulation have proved promising effects: vagus and right median nerve stimulation. Because of their extensive projections throughout the regions of the brain responsible for consciousness, the vagus and right median nerves have been targeted for studies of nerve stimulation in the treatment of coma and have proven effective in improving the states of consciousness in comatose patients. Although complete mechanisms are still unknown, many studies have corroborated the idea that delivering electrical pulses through such stimulation can elevate levels of brain activity in the stagnant coma-state brain, especially in sleep-wake cycle related brain regions such as the reticular activating system, locus coeruleus, parietal cortex, and thalamus. By activating the neurons in these parts of the brain, neural pathways — noradrenergic, cholinergic, GABAergic, dopaminergic, and orexin pathways — are enhanced. These pathways have shown to interact with each other in a complex manner still not yet thoroughly explored, but have proven to collectively increase arousal in coma patients. Furthermore, nerve stimulation seems to have positive effects on the physiological healing of the brain from traumatic brain injury, one of the root causes of coma, through increasing cerebral blood flow. Taken together, these studies point towards right median and vagus nerve stimulation as a promising treatment that can help better the outcome of coma.

Keywords: Coma, Nerve Stimulation, Vagus Nerve, Right Median Nerve, Cctivation, Reawakening

1. Introduction

A disorder of consciousness is defined as a deviation from the normal state of consciousness, typically caused by any form of damage to the brain. There are three major states of consciousness— coma, persistent vegetative state, or minimally conscious state. Firstly, the medical dictionary defines the coma state as “state of extreme unresponsiveness, in which an individual exhibits no voluntary movement or behavior”(Miller-Keane, 2003). If the patient has not regained consciousness after 2-4 weeks, the disorder

may either progress into a persistent vegetative state (PVS) or minimally conscious state (MCS), or remain in a coma state. In PVS, a patient is awake but unaware of their personal state or surroundings; in MCS, a patient is awake and shows inconsistent but clear signs of being aware.

There are multiple ways to measure one’s level of consciousness through assessing physical signs of wakefulness. In order to quantify these signs and symptoms, researchers use the Glasgow Coma Scale (GCS). This scale was developed by Professors Graham Teasdale and Bryan Jennette in 1974, and

* Corresponding Author
chloe.chan0511@gmail.com

Advisor: Kelly Bonekamp
kelly.bonekamp@gmail.com

has since been the most popular method of quantifying coma severity. The GCS tests eye opening on a scale of 1-4 points, verbal responses on a scale of 1-5 points, and motor responses on a scale of 1-6 points. An overall GCS of 8 or less is classified as severe coma state, 9-12 as moderate, and 13-15 as mild (Magee, 2021). GCS is used in a clinical setting to diagnose patients as well as used in multiple studies of coma treatment to effectively quantify improvement in consciousness following nerve stimulation.

Another means of quantifying consciousness is through the JFK Coma Recovery Scale-Revised (CRS-R), initially discovered by Giacino in 1991 (Giacino and Kalmar, 2006). This scale also assesses auditory, oromotor, communication (including visually or aurally based forms), and arousal functions in addition to the responses already monitored by the GCS (Giacino and Kalmar 2004). CRS-R is more commonly used in traumatic brain injury and such related studies, as it provides a more comprehensive view of consciousness and the improvement thereof (Giacino and Kalmar, 2006).

One of the most common causes of coma is traumatic brain injury (TBI), which is an injury to the brain caused by external forces. TBI can result in brain lesions, hemorrhage, compression and depression that may cause damage to the consciousness-related brain regions, namely the reticular activating system (RAS), locus coeruleus (LC), thalamus, and the multiple regions of the cortex. Damage to these regions may disturb the pathways which regulate wakefulness, potentially leading to brain-wide malfunctioning in the form of a coma. Speculation about motivating brain activity in these areas to restore original function are used as a basis for coma treatment studies.

The prognosis of coma largely depends on the length of coma as well as the depth of consciousness as measured by state of consciousness (PVS or MCS) and the GCS and/or CRS. The longer time in the coma state and the deeper the coma, the less likely a full recovery will be made (Faugeras, 2018). 1 out of 8 traumatic brain injury patients fall into a coma, and face a fatality rate of 30-90% depending on the presence of accompanying signs (Firsching, 2017; Lombardi, et al., 2002). And only 13% of TBI

patients with a GCS of 3 or less achieved a good functional outcome after six months (Chamoun, et al., 2009). Given the slim chances of awakening and full recovery, and the number of people with disorders of consciousness, searching for effective and safe treatment options is crucial.

2. Nerve Stimulation

Nerve stimulation has been used for hundreds of years to treat various mental disorders including depression, seizures, and now potentially, disorders of consciousness. This involves the stimulation of specific nerves to activate widespread neuron firing within the target region and its surrounding areas. When an individual is in a coma, the brain is functioning at its lowest stage of alertness, so the goal of nerve stimulation is to increase brain functioning by activating neural activity to regain consciousness (Universitat Pompeu Fabra, 2019). However, in order to ensure efficacy of nerve stimulation, specific nerves chosen for the stimulation process need to be extensive, but also selectively targeted towards the specific wakefulness-promoting regions of the brain. In addition, many studies utilize functional magnetic resonance imaging (fMRI), positron emission tomography (PET) scans, and monitoring of cerebral blood flow (CBF) as means to measure and track these changes in neural activity in different regions of the brain in order to determine improvement in nerve-stimulated consciousness.

2.1 Vagus Nerve Stimulation

Vagus nerve stimulation has been used as a reliable and tested means to treat epilepsy since its FDA approval in 1997; recent studies, however, have revealed the vagus nerve as a promising site to deliver nerve stimulation to comatose patients. The vagus nerve is the longest of the 12 cranial nerves of the body, extending from the medulla oblongata of the brainstem to the colon. The vagus nerve is also connected to many organs in the body, and is important in carrying sensory information from these organs to the brain. In order to carry out this important function, the vagus nerve has extensive

projections from the medulla oblongata in the hindbrain to consciousness- regulating brain regions in the mid and forebrain areas. This makes the vagus nerve a strong candidate as stimulation of this nerve may activate multiple cortices of the brain that help to contribute towards increased arousal.

Vagus nerve stimulation is usually initiated via a minimally invasive procedure to implant a device. Firstly, a small incision is made on the left side of the chest for the pulse generator and on the collarbone where the thin wires and electrodes are placed to connect the generator to the vagus nerve, thus allowing for electrical impulses to be sent to the nerve. Nerve stimulation intensity and intervals are determined by the programming of the pulse generator, and are usually determined by doctors monitoring the patient.

In the first reported case study of vagus nerve stimulation, a 73 year old female patient remained in a VS for 50 days following respiratory and cardiac arrest and unsuccessful cardiopulmonary resuscitation. Three doctors who were not a part of the study assessed the patient using the CRS-R and the average of the two closest scores were taken to determine the starting level of consciousness. VNS was then implemented twice per day, lasting 20 minutes each session, for a total duration of one month. fMRI revealed activation of the thalamus and posterior cingulate precuneus, brain structures responsible for memory and visuospatial processing, following vagus nerve stimulation. Increased connectivity between the aforementioned regions and the ventral medial prefrontal cortex and superior temporal gyrus was also observed post-VNS, which may indicate increased cognitive and auditory processing, respectively. At the end of the treatment, although complete consciousness was not restored, the patient showed promising signs of increased wakefulness, including a diagnosis from vegetative to the minimally conscious state and an increase from 6 to 13 on the CRS-R (Yu, et al., 2017).

The results of this case study are supported by another study conducted on in vivo mouse models. This study used calcium imaging to measure the activity of neurons across the cortex of the brain following vagus nerve stimulation. Stimulation of the vagus nerve caused widespread activation of neurons

in visuospatial processing regions like the medial to lateral somatosensory, motor and retrosplenial cortical regions (Collins, et al., 2021). Responses in these regions of the brain were found to be positively correlated with signs of wakefulness, such as whisker movement. Taken together, these results suggest that vagus nerve stimulation can be used to increase brain activity in the sensorimotor cortices of the brain, as well as improve both clinical and physical signs of arousal in coma patients. Further studies corroborate these findings, and even suggest that vagus nerve stimulation can also facilitate healing from traumatic brain injury (Shi, et al., 2013).

In a more recent study, patients above 17 years old diagnosed with severe TBI were given VNS up to 0.5 mA for an eight week period for up to 4 hours daily to increase wakefulness. 60% of the patients in the study exhibited over a 3 point improvement in the CRS-R score. As seen in past studies, there were no adverse side effects present in any of the patients, other than mild skin irritation of the stimulation spot. (Hakon, et al., 2020; Kreuzer, et al., 2012) Preclinical trials have assessed and confirmed cardiac safety when performing nerve stimulation (Kreuzer, et al., 2012), and multiple studies also support the safety of VNS usage (Redgrave, et al., 2018). With that being said, the efficacy of VNS in treating comatose patients are still being tested today in hopes of its implementation into coma treatment (Vitello, et al., 2023; Noe, et al., 2019).

2.2 Right Median Nerve Stimulation

The right median nerve is yet another potential target for stimulation in the treatment of coma. This nerve is both a sensory and motor nerve which stems from the lateral and medial cords of the brachial plexus, a network of nerves in the neck, and runs through the upper and lower spinal cord before finishing at the hands and fingers (Cleveland Clinic, 2021). Stimulation typically occurs via the placement of electrodes in the forearm. These electrodes are then controlled by a device which sends electrical pulses of different amplitudes and intensities through the electrodes to stimulate right median nerve connected structures (Lei, et al., 2015).

The first study utilizing the right median nerve to

hasten awakening from coma was conducted by researcher J. Bryan Cooper of the East Carolina University School of Medicine in 1998. The right median nerve was targeted because of clinical observations of increased alertness and increased activity in the Broca's area (an area important for speech production) following right median nerve stimulation. A group of 6 patients, each having been in a coma for at least 3 months due to traumatic brain injury, were admitted to the study. Patients were stabilized, and given a score of 4-8 on the Glasgow Coma Scale prior to right median nerve stimulation. Patients were then randomly assigned to either control or treatment group; the procedure was double blind in order to minimize bias. Rubber electrodes were then placed on the experimental patients' right distal forearm as a direct pathway to the brain via the right median nerve. 8 or 12 hours of stimulation was implemented daily for 2 weeks. At the end of the stimulation period, measurement of consciousness was quantified by two major factors: GCS scores and days spent in the intensive care unit (ICU). Patients who received right median stimulation had an average GCS improvement of 4.0 while patients who did not receive right median stimulation had an average of 0.7, indicating a significant difference and providing evidence of right median nerve stimulation's effect in increasing arousal. ICU results further confirmed the conclusions from the GCS scores; patients who received stimulation treatment stayed in the ICU on average 10 days less than patients in the control group (Cooper, et al., 1999).

Much like the vagus nerve, the right median nerve shows connectivity with the reticular activating system (a sleep-wake regulating structure in the brain), and therefore may be an effective target for nerve stimulation to hasten awakening from coma (Peri, et al., 2001).

More comprehensive studies have started looking into the means behind right median nerve stimulation's effectiveness. A concept of measuring cerebral blood flow (CBF), which indicates potential enhancement of metabolism and brain activity throughout the brain, was applied to and measured in right median nerve stimulation patients before, throughout and after stimulation. It was seen that in patients who received right median nerve stimulation,

there was a significant increase in CBF compared to patients in the control group (Liu, et al., 2003), suggesting right median nerve stimulation's overall effectiveness in increasing brain activity.

Another study provided evidence of which specific regions were activated by such stimulation. In the study, 10 healthy male patients underwent right median nerve stimulation at multiple frequencies, ranging from 0.2 to 20 Hz. PET scans were taken every 10 minutes to monitor CBF as a metric for changes in neural activity in different regions of the brain. It was found that there was a significant increase in CBF in the sensorimotor and parietal cortices at a frequency of 4 Hz. Nearby sulci and gyri, namely the central sulcus (a brain structure important in transporting blood the lateral brain that allows it to function) and the sylvian fissure also saw an increase in CBF (Ibanez, et al., 1995; Griffiths, et al., 2010). These results are significant because they identify regions of the brain activated by stimulation, help researchers form hypotheses of the mechanisms behind stimulation, and suggest potential pathways to target in the treatment of coma.

Throughout multiple studies, traumatic brain injury patients in the comatose state have responded positively to right median nerve stimulation with increase in Glasgow Coma and JFK-CRS-R scores, (Lei, et al., 2015; Sharma, et al., 2015; Straughn and Denais, 2019), physical signs of arousal (Liu, et al., 2003), and in some cases even the regaining consciousness (Liu, et al., 2003; Lei, et al., 2015). A more recent study examined the effect of right median nerve stimulation on comatose patients suffering from heavy TBI. These patients received two weeks of nerve stimulation along with standard treatment which resulted in increased brain functioning and early awakening in these patients (Jia, et al., 2022).

3. Mechanisms of Action

As both vagus nerve and right median nerve stimulation target similar areas of the brainstem and forebrain such as the reticular activating system, locus coeruleus, thalamus, and cortex; both use and produce similar effects at a neural activity level. Many of these different neurotransmitter pathways

overlap with one another (due to their proximity within the aforementioned target regions) in a complex manner that is still not completely understood today. However, studies have slowly begun to reveal the potential mechanisms of action with how these pathways work to regulate consciousness, and the effects of nerve stimulation in modulating these pathways.

3.1 Norepinephrine

Norepinephrine has long been associated with increased wakefulness and attention (Larner, 2002). Therefore, both vagus and right median nerve stimulation are able to regulate and increase consciousness by producing an acute increase in norepinephrine levels in the cortical and hippocampal regions of the brain (Follesa, et al., 2007; Lei, et al., 2015; Kayama and Koyama, 1998). This response may be explained by the vagus nerve's projections to the nucleus tractus solitarii (NTS) in the medulla, which then projects to other areas including the locus coeruleus, the main production center of norepinephrine (Cooper, et al., 1999; Berger, et al., 2021). Once the locus coeruleus is activated by nerve stimulation, norepinephrine is released from this small but extensive brain structure to the rest of the brain and spinal cord, thus helping to increase overall wakefulness (Schwarz and Luo, 2015).

3.2 Acetylcholine

Additional studies show how exactly this happens via interactions between norepinephrine and acetylcholine: norepinephrine from the locus coeruleus is released through its projections and onto the basal forebrain. As norepinephrine is released, it binds to the adrenergic receptors on the cholinergic neurons and activates them. Activation of these cholinergic neurons occurs when the locus coeruleus neurons are active in order to activate cholinergic cells, which occurs during periods of waking (Jones, 2004). So when these norepinephrine-activated cholinergic cells then rapidly fire, they release acetylcholine from the basal forebrain. This release allows for information exchange between the

thalamus and the cortex, and induces widespread cortical activity to increase wakefulness (Vazquez and Baghdoyan, 2001; Cooper, et al., 1999). Note that both of these studies demonstrate right median nerve stimulated increase in activity within the thalamocortical regions of the brain which have been correlated with increased consciousness (Steriade, 1981), while inactivation in these corresponding areas has promoted the onset of sleep (Magnin, et al., 2010). Right median and vagus nerve stimulation induces a norepinephrine-modulated acetylcholine release from the reticular activating system, which may play a key role in regulation of activity in the thalamus and its projections into the cortical regions, promoting overall wakefulness and sensorimotor processing.

3.3 Orexin

Another neural pathway involves orexin and orexin-producing neurons located in the lateral hypothalamic area, which interact with monoamine pathways present in regions of the brain activated via nerve stimulation where important sleep-wake cycle regulating regions reside (Sakurai, 2007).

More specifically, there are two types of orexin neuropeptides, orexin-A and orexin-B, both of which are released in the lateral hypothalamus and once bind to their receptors, orexin receptor 1 / orexin receptor 2 (OXR1/2), promotes wakefulness possibly by its connections to monoaminergic systems as previously mentioned. Because the vagus nerve connects indirectly to the lateral hypothalamus via the nucleus tractus solitarii and parabrachial nucleus (Barry and Eleni, 2022), studies have hypothesized and shown that stimulation of the vagus nerve would cause an upregulation of orexin-A to bind to OXR1 within those parts of the brain to promote wakefulness (Wang, et al., 2018; Dong and Feng, 2018; Dong and Papa, 2018). More comprehensive results show that 1) when OXR1 receptors are blocked with an antagonist, rats tend to wake up less from comas despite vagus nerve stimulation, and 2) orexin-A is upregulated immediately following traumatic brain injury and loss of consciousness (Wang, et al., 2018).

Further research demonstrates how right median nerve / median nerve stimulation promotes arousal from coma by these same orexin mechanisms (Chen, et al., 2018), potentially through stimulation of regions of the brain namely the reticular activating system and intralaminar nuclei of the thalamus (Zhong, et al., 2015; Feng and Du, 2016). These findings indicate the importance of the orexin-A and OXR1 in the brain's natural attempts at reawakening, and nerve stimulation's role in upregulation of these neuropeptides. There have been additional studies that have shown, however, that OXR2 and dual orexin receptor antagonists, and not only OXR1 antagonists alone, inhibit wakefulness; this suggests the importance of both neuropeptides and their receptors in increasing consciousness (Kalogiannis, et al., 2011).

Some combination of OX1R and OX2R mRNA, which codes for the different orexin receptors, were expressed in cholinergic, noradrenergic, GABAergic and serotonergic neurons, indicating orexin importance in regulation of these monoamine neural activity (Mieda, et al., 2011), and suggesting the possibility of maintaining consciousness through modulation of orexin expression/orexin release to regulate wakefulness related monoaminergic/ cholinergic pathways (Mieda, et al., 2004).

3.4 GABAergic Pathways

There has also been speculation regarding the interconnectivity between cholinergic and GABAergic neurons in the basal forebrain in which both are located near each other and stimulation of cholinergic neurons also caused firing of GABAergic neurons (Yang, et al., 2014). Some studies hypothesize that cholinergic and related monoamine pathways inhibit GABAergic neurons to promote wakefulness (Jones, 2004). Still, more recent studies corroborate the opposite, that these cholinergic and noradrenergic pathways increase GABA function to produce the same effects (Zank, et al., 2016); although both studies seem contradictory, their results do suggest the same hypothesis that GABA neurons play a more extensive role in regulating consciousness than previously thought to have

played, debunking the idea that its only function is the inhibition of all other neural activity. However, a more recent study may suggest a way to corroborate both these findings: depending on where GABAergic neurons are active, this may either promote or inhibit wakefulness. Through a combination of viral tracing, circuit mapping, electrophysical recordings, and optogenetics in mice, researchers found that GABAergic neurons in the lateral hypothalamus selectively inhibit certain GABAergic neurons whilst allowing a large portion of the uninhibited neurons in the dorsal raphe nucleus fire, thereby promoting arousal (Gazea, et al., 2021; Cai, et al., 2022). The opposite is also true: an inhibition of GABAergic neurons in the hypothalamus would result in its inability to interact with GABAergic neurons in the dorsal raphe nucleus and a reduction in wakefulness (Huang, et al., 2020).

3.5 Dopamine and Serotonin Parallel Functioning

Additionally, another neurotransmitter known as dopamine functions similarly to norepinephrine in maintaining wakefulness, arousal, and attention. Consistent overlaps in the two neural pathways have led researchers to hypothesize the parallel functioning between the two neurotransmitters, as dopamine is also released from the locus coeruleus as a cotransmitter to norepinephrine (Slamloo and Fazlali, 2020). Further research has correlated right median nerve stimulation with increased levels of dopamine that contribute to an increase in arousal in comatose patients (Jia, et al., 2022).

In another study, serotonin is demonstrated to work in opposition with dopamine and norepinephrine functions. Mutant mice that don't express the serotonin gene for the serotonin receptor (meaning serotonin responses were inhibited) were demonstrated to have a greater amount of REM sleep and decreased slow wave sleep. These signs of increased consciousness and decreased serotonin release may be related to increased release of norepinephrine and dopamine (Monti, 2010). Additional electrophysiological data show that serotonin can exert inhibitory effects on dopamine – parallel to increased levels of dopamine that coupled

with a decreased serotonin response – both directly and potentially indirectly through modification of GABAergic signals (Giovanni, et al., 2008). Again, these are loosely related observations that suggest a potential in targeting different neurotransmitters through drug therapy in treatment of coma.

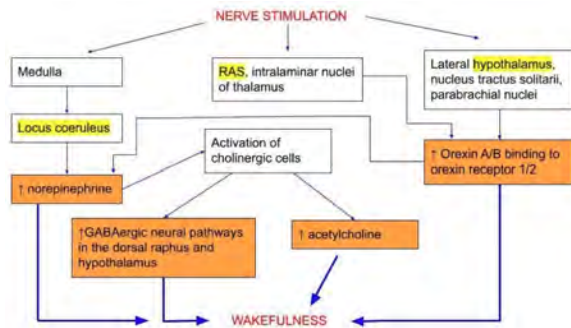


Figure 1. A physical representation of the multiple pathways activated by nerve stimulation that lead to wakefulness. Highlighted are brain regions crucial in the regulation of consciousness, and the orange boxes are the distinct pathways aforementioned.

4. Physiological Responses to Nerve Stimulation

Nerve stimulation also facilitates physiological responses which aid the physical healing of the brain from traumatic brain injury, the root cause of coma and disorders of consciousness. Some studies suggest that this healing is due to increases in cerebral blood flow immediately following vagus / right median nerve stimulation to different areas of the brain (Ibanez, et al., 1995; Liu, et al., 2003; Cooper, et al., 1999; Schiff, 2008; Henry, et al., 1998). As the blood delivers glucose to the brain, that glucose is metabolized by neurons for energy and used to send neural impulses, which may aid increased wakefulness and eventual recovery from coma. One such neural pathway activated by increased cerebral blood flow is the initiation of release of acetylcholine via vagus nerve stimulation, which inhibits an inflammatory response in the brain and promotes arousal from coma (Shi, et al., 2013; Collins, et al., 2021). This anti-inflammatory effect of vagus nerve stimulation as shown in multiple studies (Neren, et al., 2015; Bonaz, et al., 2013; Liu, et al., 2020) has been used in treatment of traumatic brain injury,

which makes using it in the treatment of coma the logical subsequent step. Lowering swelling in the brain via nerve stimulation would allow for the brain to heal and regain functioning consciousness.

5. Discussion

Disorders of consciousness are highly common and devastating conditions for both the patients and their families. However, at the moment, there is still no approved definitive treatment for a comatose patient; the general protocol is to try to reverse the coma and reawaken the patient using the help of medication (MayoClinic 2020). Recent studies have shown potential, however, in the treatment of coma via nerve stimulation to deliver pulses to activate regions of the brain connected to the selectively targeted right median and vagal nerves. Mechanisms include general increased brain metabolism, connectivity and functioning via neural pathways (mainly noradrenergic and cholinergic pathways), all of which interconnect with each other, though exact interactions between these mechanisms and how they work together to promote consciousness are still unclear.

With nerve stimulation's potential, also comes more factors that should be thoroughly tested and resolved before its approval in a clinical setting. The first factor being the actual efficacy in its treatment of all comas; the longer a patient is in a coma, generally the more severe the coma is and therefore harder it is to stimulate reawakening (Sherer, et al., 2008). In many of these studies, patients have been in comas for at least several months and although subjects have demonstrated signs of increased consciousness and brain activity, a percentage of coma patients within these studies still have yet to make a complete recovery in regaining consciousness. This suggests the potential in longer stimulation periods to aide in the full recovery of coma; however, that in and of itself also entails complications; long term nerve stimulation, especially in vagus nerve stimulation where the stimulator is installed surgically, causes side effects including nerve damage, cognitive side effects, paresthesias, etc. (Ben-Menachem, 2001). Furthermore, although nerve stimulation may help trigger signs of consciousness with TBI-induced

comas, it has not proved as effective when treating coma caused by other factors such as hypoxia (Liu, et al., 2003). However, these side effects are rare and only present in extreme cases. When used in the short term, multiple studies have come to the consensus that vagus and right median nerve stimulation are safe, low-cost, and an effective means to increase wakefulness in comatose patients (Straughn and Denais, 2019; Hakon, et al., 2020)

Researchers continue to explore the safety and efficacy of right median and vagus nerve stimulation for recovery of consciousness on a wider level. One of which the ACES Trial studying right median nerve stimulation in which 360 patients from the neurocenters of China, India, Nepal and Kazakhstan will undergo nerve stimulation treatment. Results have still yet to be published, but if successful will provide the evidence needed to implement its widespread usage in hospitals across Asia and potentially across the globe (ACES, 2016). Similar trials on a smaller scale are being conducted for vagus nerve stimulation (Vitello, et al., 2023; Noe, et al., 2019). Studies exploring the potential of nerve stimulation as a whole in the treatment of coma have continued, including other stimulation types such as spinal cord stimulation and deep brain stimulation (Bai, et al., 2017; Rezaei Haddad, et al., 2019). Nerve stimulation has the potential to activate or inhibit these complex systems in order to promote arousal and even reawakening from coma.

Acknowledgement

I would like to acknowledge and thank my mentor Kelly Bonekamp for her knowledge and expertise in the process of reviewing previous studies, and her guidance through writing, revising, and publishing my paper. I would also like to thank Polygence for the opportunity to research, and my parents for helping me to achieve this publishing.

References

American Association of Neurological Surgeons. (n.d.). Vagus nerve stimulation. <https://www.aans.org/en/Patients/Neurosurgical-Conditions-and-Treatments/Vagus-Nerve-Stimulation>

Asia Coma Electrical Stimulation (the ACES Trial) (ACES). (2016, January 1). NIH U.S. National Library of Medicine — ClinicalTrials.gov. Retrieved December 18, 2022, from <https://clinicaltrials.gov/ct2/show/study/NCT02645578>.

Awareness. (2003). Retrieved April 28th, 2020, from <https://medical-dictionary.thefreedictionary.com/awareness>

Bai Y., et al. (2017, March) Spinal cord stimulation modulates frontal delta and gamma in patients of minimally consciousness state. *Neuroscience*. <https://doi.org/10.1016/j.neuroscience.2017.01.036>.

Ben-Menachem, E. (2001). Vagus nerve stimulation, side effects, and long term safety. *PubMed*. <https://doi.org/10.1097/00004691-200109000-00005>

Berger, A., et al. (2021, December 2). How is the norepinephrine system involved in the antiepileptic effects of vagus nerve stimulation? <https://doi.org/10.3389/fnins.2021.790943>

Bonaz, B., et al. (2013) Vagus nerve stimulation: from epilepsy to the cholinergic anti-inflammatory pathway. *Neurogastroenterology & Motility*, 25(3), 208-221. <https://doi.org/10.1111/nmo.12076>

Cai, P., et al. (2022) Regulation of wakefulness by GABAergic dorsal raphe nucleus-ventral tegmental area pathway. *Sleep*, 45, 235. <https://doi.org/10.1093/sleep/zsac235>

Chamoun, R.B., Robertson, C.S., & Gopinath, S.P. (2009) Outcome in patients with blunt head trauma and a Glasgow Coma Scale score of 3 at presentation. *J Neurosurg*. <https://doi.org/10.3171/2009.2.JNS08817>

Chen, Y. H., et al. (2018) Median nerve stimulation induces analgesia via orexin-initiated endocannabinoid disinhibition in the periaqueductal gray. <https://doi.org/10.1073/pnas.1807991115>

Cleveland Clinic. (2021, September 24). Median nerve. Cleveland Clinic. <https://my.clevelandclinic.org/health/body/21889-median-nerve>

Collins, L., et al. (2021). Vagus nerve stimulation

induces widespread cortical and behavioral activation. *Current Biology*.
<https://doi.org/10.1016/j.cub.2021.02.049>.

Coma. (2020, November 20). Retrieved April 22, 2022, from
<https://www.mayoclinic.org/diseases-conditions/coma/diagnosis-treatment/drc-20371103>

Cooper, J. B., et al. (1999, April). Right median nerve electrical stimulation to hasten awakening from coma. *PubMed*.
<https://doi.org/10.1080/026990599121638>

Dong, X. Y., et al. (2018). Vagus nerve stimulation causes wake-promotion by affecting neurotransmitters via orexins pathway in traumatic brain injury induced comatose rats. *International Journal of Clinical and Experimental Medicine*.
<https://e-century.us/files/ijcem/11/5/ijcem0067767.pdf>

Dong, X.Y. & Feng Z. (2018, February) Wake-promoting effects of vagus nerve stimulation after traumatic brain injury: upregulation of orexin-A and orexin receptor type 1 expression in the prefrontal cortex. *Neural Regen Res*.
<https://doi.org/10.4103/1673-5374.226395>.

Faugeras, F., et al. (2018) Survival and consciousness recovery are better in the minimally conscious state than in the vegetative state, *Brain Injury*, 32:1, 72-77,
<https://doi.org/10.1080/02699052.2017.1364421>

Feng, Z., & Du, Q. (2016, June). Mechanisms responsible for the effect of median nerve electrical stimulation in traumatic brain injury-induced coma: Orexin-A-mediated N-methyl-D-aspartate receptor subunit NR1 upregulation. *PubMed*.
<https://doi.org/10.4103/1673-5374.184494>

Firsching, R. (2017, May 5). Coma after acute head injury. *PubMed*.
<https://doi.org/10.3238/arztebl.2017.0313>

Follesa, P., et al. (2007). Vagus nerve stimulation increases norepinephrine concentration and the gene expression of BDNF and bFGF in the rat brain. *PubMed*.
<https://doi.org/10.1016/j.brainres.2007.08.045>

Gazea, M., et al. (2021) Reciprocal Lateral

Hypothalamic and Raphe GABAergic Projections Promote Wakefulness.
Journal of Neuroscience 2, 41(22), 4840-4849.
<https://doi.org/10.1523/JNEUROSCI.2850-20.2021>

Giacino, J & Kalmar, K. (2006). Introduction to the JFK Coma Recovery Scale-Revised (CRS-R). The Center for Outcome Measurement in Brain Injury.
<http://www.tbims.org/combi/crs>

Giacino, J. T., & Kalmar, K. (2004). CRS-R coma recovery scale-revised - administration and scoring guidelines. SRALAB.
<https://www.sralab.org/sites/default/files/2017-06/coma%20recovery%20scale%20-%20revised%20%28crs-r%29.pdf>

Giuseppe, G. D., et al. (2008) Serotonin-dopamine interaction: electrophysiological evidence. In Giuseppe, G. D., et al. (Eds.), *Progress in Brain Research*. (pp. 45-71).
[https://doi.org/10.1016/S0079-6123\(08\)00903-5](https://doi.org/10.1016/S0079-6123(08)00903-5).

Griffiths, P. D., et al. (2010) Section 1 - Surface Anatomy of the Brain. *Atlas of Fetal and Postnatal Brain MR* (Pages 7-34)
<https://doi.org/10.1016/B978-0-323-05296-2.50004-7>

Hakon, J., et al. (2020) Transcutaneous Vagus Nerve Stimulation in Patients With Severe Traumatic Brain Injury: A Feasibility Trial. *Neuromodulation: Technology at the Neural Interface*, 23, 721-892.
<https://doi.org/10.1111/ner.13148>

Henry T. R., et al. (1998, September) Brain blood flow alterations induced by therapeutic vagus nerve stimulation in partial epilepsy: I. Acute effects at high and low levels of stimulation. *Epilepsia*.
<https://doi.org/10.1111/j.1528-1157.1998.tb01448.x>.

Huang, Y., Li, Y., Leng, Z. W. (2020) Melatonin inhibits GABAergic neurons in the hypothalamus consistent with a reduction in wakefulness. *NeuroReport*, 31(2), 92-98.
<https://doi.org/10.1097/WNR.0000000000001374>

Ibáñez, V., et al. (1995, October) Effects of stimulus rate on regional cerebral blood flow after median nerve stimulation. *Oxford Academic*.
<https://doi.org/10.1093/brain/118.5.1339>

- Jia, Y., et al. (2022) MicroRNA alteration in cerebrospinal fluid from comatose patients with traumatic brain injury after right median nerve stimulation. <https://doi.org/10.1007/s00221-022-06414-7>
- Jones, B. E. (2004). Activity, modulation and role of basal forebrain cholinergic neurons innervating the cerebral cortex. In L. Descarries, K. Krnjevic, M. Steriade (Eds.), *Acetylcholine in the Cerebral Cortex* (pp. 157–169). ScienceDirect. [https://doi.org/10.1016/S0079-6123\(03\)45011-5](https://doi.org/10.1016/S0079-6123(03)45011-5)
- Jones, B. E. (2004). Activity, modulation and role of basal forebrain cholinergic neurons innervating the cerebral cortex. PubMed. [https://doi.org/10.1016/S0079-6123\(03\)45011-5](https://doi.org/10.1016/S0079-6123(03)45011-5)
- Kalogiannis, M., et al. (2011, April 13). Cholinergic modulation of narcoleptic attacks in double orexin receptor knockout mice. *PLOS One Journals*. <https://doi.org/10.1371/journal.pone.0018697>
- Kayama, Y., & Koyama, Y. (1998). *Brainstem Neural Mechanisms of Sleep and Wakefulness*. Karger. <https://doi.org/10.1159/000052235>
- Komisaruk B. R., & Frangos E. (2022) Vagus nerve afferent stimulation: Projection into the brain, reflexive physiological, perceptual, and behavioral responses, and clinical relevance. *Autonomic Neuroscience*. <https://doi.org/10.1016/j.autneu.2021.102908>.
- Kreuzer, P. M., et al. (2012) Transcutaneous vagus nerve stimulation: retrospective assessment of cardiac safety in a pilot study. <https://doi.org/10.3389/fpsy.2012.00070>
- Larner A. J. (2002, November) *Neurochemistry of consciousness: Neurotransmitters in mind*. (Brain, Volume 125, Issue 11, Pages 2581–2582). <https://doi.org/10.1093/brain/awf244>
- Lei, J., et al. (2015, October 15). Right median nerve electrical stimulation for acute traumatic coma patients. PubMed. <https://doi.org/10.1089/neu.2014.3768> <https://pubmed.ncbi.nlm.nih.gov/25664378/>
- Levy, D. E. et al. (1981, March). Prognosis in nontraumatic coma. PubMed. <https://doi.org/10.7326/0003-4819-94-3-293>
- Liu, C. H., et al. (2020) Neural networks and the anti-inflammatory effect of transcutaneous auricular vagus nerve stimulation in depression. *Journal of Neuroinflammation*, 17, 54. <https://doi.org/10.1186/s12974-020-01732-5>
- Liu, J. T., et al. (2003). Regaining consciousness for prolonged comatose patients with right median nerve stimulation. *Neurosurgical Re-Engineering of the Damaged Brain and Spinal Cord*. https://doi.org/10.1007/978-3-7091-6081-7_3
- Lombardi, F., et al. (2002, April 22). Sensory stimulation for brain injured individuals in coma or vegetative state. PubMed. <https://doi.org/10.1002/14651858.cd001427>
- Magee, D. J. (n.d.). Head and face. In J. L. McDougal (Eds.), *Orthopedic physical assessment* (pp. 73-163). ScienceDirect. <https://doi.org/10.1016/B978-0-323-05226-9.50011-6>
- Magnin M., et al. (2010, February) Thalamic deactivation at sleep onset precedes that of the cerebral cortex in humans. *Proc Natl Acad Sci USA*. <https://doi.org/10.1073/pnas.0909710107>.
- Mieda M., et al. (2004, March) Orexin peptides prevent cataplexy and improve wakefulness in an orexin neuron- ablated model of narcolepsy in mice. *Proc Natl Acad Sci USA*. <https://doi.org/10.1073/pnas.0400590101>.
- Mieda, M., et al. (2011, April 27). Differential roles of orexin receptor-1 and -2 in the regulation of non-REM and REM sleep. *Journal of Neuroscience*. <https://doi.org/10.1523/JNEUROSCI.6506-10.2011>
- Monti, J. M. (2010, October) The role of dorsal raphe nucleus serotonergic and non-serotonergic neurons, and of their receptors, in regulating waking and rapid eye movement (REM) sleep. ScienceDirect. <https://doi.org/10.1016/j.smr.2009.10.003>.
- National Health Services. (2018, August 6). Overview disorders of consciousness. NHS UK. <https://www.nhs.uk/conditions/disorders-of-consciousness/>
- Neren, D., et al. (2015) Vagus nerve stimulation and other neuromodulation methods for treatment of traumatic brain injury. *Neurocritical Care*, 24, 308-319. <https://doi.org/10.1007/s12028-015-0203-0>

- Noe, E., et al. (2019) Feasibility, safety and efficacy of transauricular vagus nerve stimulation in a cohort of patients with disorders of consciousness. *Brain Stimulation*, 13, 277-522. <https://doi.org/10.1016/j.brs.2019.12.005>
- Peri C. V., et al. (2001, October) Pilot study of electrical stimulation on median nerve in comatose severe brain injured patients: 3-month outcome. *Brain Inj.* <https://doi.org/10.1080/02699050110065709>.
- Redgrave, J., et al. (2018) Safety and tolerability of Transcutaneous Vagus Nerve stimulation in humans; a systematic review. *Brain Stimulation*, 11, 1225-1238. <https://doi.org/10.1016/j.brs.2018.08.010>
- Rezaei Haddad, A., Lythe V., & Green A. L. (2019, June) Deep Brain Stimulation for Recovery of Consciousness in Minimally Conscious Patients After Traumatic Brain Injury: A Systematic Review. *Neuromodulation*. <https://doi.org/10.1111/ner.12944>
- Sakurai T. (2007, March) The neural circuit of orexin (hypocretin): maintaining sleep and wakefulness. *Nat Rev Neurosci*. <https://doi.org/10.1038/nrn2092>.
- Schiff N. D. (2008) Central thalamic contributions to arousal regulation and neurological disorders of consciousness. *Ann NY Acad Sci*. <https://doi.org/10.1196/annals.1417.029>.
- Schwarz, L. A., & Luo, L. (2015, November 2). Organization of the locus coeruleus-norepinephrine system. *Current Biology*. <https://doi.org/10.1016/j.cub.2015.09.039>
- Sharma, V., Kaur, H., & Gupta, D. (2015). Right median nerve stimulation for improving consciousness: A case series. <https://doi.org/10.1055/s-0035-1569473>
- Sherer M., et al. (2008, June) Comparison of indices of traumatic brain injury severity: Glasgow Coma Scale, length of coma and post-traumatic amnesia. *J Neurol Neurosurg Psychiatry*. <https://doi.org/10.1136/jnnp.2006.111187>
- Shi, C., Flanagan, S. R., & Samandani, U. (2013, November 12). Vagus nerve stimulation to augment recovery from severe traumatic brain injury impeding consciousness: A prospective pilot clinical trial. *Neurological Research*. <https://doi.org/10.1179/1743132813Y.000000016>
- Slamloo, Y. R., & Fazlali, Z. (2020, January 21). Dopamine and noradrenaline in the brain; Overlapping or dissociate functions? *Frontiersin*. <https://doi.org/10.3389/fnmol.2019.00334>
- Steriade, M. (2010, February 4). EEG desynchronization is associated with cellular events that are prerequisites for active behavioral states. Cambridge University Press. <https://doi.org/10.1017/S0140525X00010037>
- Straughn, J., & Denais, Z. (2019, October 1). Application of right median nerve stimulation in an inpatient rehabilitation setting for disorders of consciousness. *Archives of physical medicine and rehabilitation*. <https://doi.org/10.1016/j.apmr.2019.08.115>
- Universitat Pompeu Fabra - Barcelona. (2019) Consciousness rests on the brain's ability to sustain rich dynamics of neural activity. *ScienceDaily*. www.sciencedaily.com/releases/2019/02/190212104217.htm.
- University of Glasgow. (2014, July 14). 40 years of the glasgow coma scale. Gla.ac.uk. https://www.gla.ac.uk/news/archiveofnews/2014/july/headline_346523_en.html
- Vazquez J., & Baghdoyan H. A. (2001, February) Basal forebrain acetylcholine release during REM sleep is significantly greater than during waking. *American Journal of Physiology*. <https://doi.org/10.1152/ajpregu.2001.280.2.R598>
- Vitello, M. M., et al. (2023) Transcutaneous vagal nerve stimulation to treat disorders of consciousness: Protocol for a double-blind randomized controlled trial. *International Journal of Clinical and Health Psychology*, 23, 100360. <https://doi.org/10.1016/j.ijchp.2022.100360>
- Wang, C., et al. (2018, June 28). The orexin/receptor system: Molecular mechanism and therapeutic potential for neurological diseases. *Frontiersin*. <https://doi.org/10.3389/fnmol.2018.00220>
- Wu, X., et al. (2017, July 10). Right median nerve electrical stimulation for acute traumatic coma (the asia coma electrical stimulation trial): Study protocol for randomised controlled trial. *PubMed*.

<https://doi.org/10.1186%2Fs13063-017-2045-x>

Yang, C., et al. (2014, February 19). Cholinergic neurons excite cortically projecting basal forebrain gabaergic neurons. *The Journal of Neuroscience*. <https://doi.org/10.1523/JNEUROSCI.3235-13.2014>

Yu, Y.T., et al. (2017) Transcutaneous auricular vagus nerve stimulation in disorders of consciousness monitored by fMRI: The first case report. *Brain Stimul.* <https://doi.org/10.1016/j.brs.2016.12.004>.

Zank, J. C., et al. (2016, February 10). Cholinergic neurons in the basal forebrain promote wakefulness by actions on neighboring non-cholinergic neurons: An opto-dialysis study. *The Journal of Neuroscience*. <https://doi.org/10.1523%2FJNEUROSCI.3318-15.2016>

Zhong, Y. J., et al. (2015). Wake-promoting actions of median nerve stimulation in TBI-induced coma: An investigation of orexin-A and orexin receptor 1 in the hypothalamic region. *Molecular Medicine Reports*. <https://doi.org/10.3892/mmr.2015.3898>

Effect and Significance of Temperature on the Speed of Electromagnetic Waves

Una Ding^{1*}

¹Uppingham School, Uppingham, Rutland, United Kingdom

Received May 22, 2022; Revised December 9, 2022; Accepted, December 28, 2022

Abstract

This project investigates the effect of temperature on the speed of electromagnetic wave when passing through water. In this experiment, a laser beam is fired through a semi-circular refraction dish containing water at an angle where the beam total internally reflected. After measuring the critical angle, calculations were made to derive the refractive index of water and ultimately the laser speed. The refractive index of water and hence speed of electromagnetic wave could be calculated. At the end of the investigation, there are no observable changes in the critical angle. Further research revealed that as the temperature increases, the speed of electromagnetic wave increases. However this change is negligible as the temperature range investigated is narrow. This effect's cause is a slight decrease in liquid density, which decreases the refractive index of water when it is heated as a result.

Keywords: Electromagnetic Wave, Total Internal Reflection, Critical Angle, Density, Refractive Index, Superposition, Wave-Particle Duality

1. Introduction

Life on Earth is dependent on water. It is the most important biological liquid in organisms. Water is the main component in human body fluids such as: interstitial fluid, blood and saliva, water is the main component. Therefore, knowing whether temperature affects the speed of electromagnetic waves traveling through water is essential to biomedical optics.

The observations made from the laser's change in critical angle is also a product of the changing refractive index of liquid. This promotes consideration of bringing temperature as a factor into fine calculations in refractometry, the study of refractive index of a material. Refractometry is widely applied in modern industry for checking how pure and concentrated liquids are; Checking the level of ripeness of fresh fruit in the food industry; checking the specific gravity of protein and urine in

plasma of animals for veterinary (Mettler-Toledo International Inc, 2021). Optical scientists need to know whether their results are affected by temperature, and whether they need to perform their tests in a temperature-controlled environment.

This investigation aims to determine the significance of a change in the speed of electromagnetic wave under the effect of temperature. Following the hypothesis of the investigation as temperature increases, the speed of electromagnetic wave should decrease, as photons will be interfered by vibrating particles in liquid. That is, increasing in temperature increases collisions between photons and water molecules, decreasing its speed.

2. Materials and Methods

* Corresponding Author
unading0@gmail.com

Advisor: Emma Ellis
ele@uppingham.co.uk

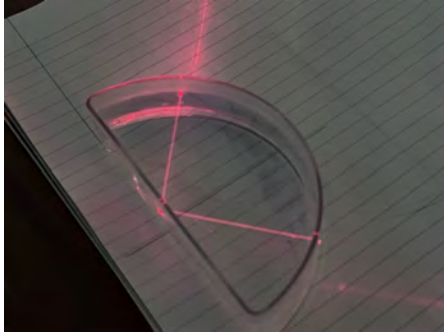


Figure 1, total internal reflection of laser passing through refraction dish filled with water.

To begin with, the shape of the refraction dish is traced with a pencil onto the middle of a blank A4 paper. The normal to the semicircle is drawn. The A4 sheet is then placed below the refraction dish, where the trace completely overlaps it. Then, water is heated in a kettle to 100°C and poured into the refraction dish. In figure 1, the first set up of this project, a monochromatic red laser is fired toward the normal at an angle adjusted where it achieved total internally reflection. A digital thermometer is used immediately after the total internal reflection to measure the current temperature of the water. The position where the laser entered the dish is marked on the A4 sheet. The dish is then removed and an incident ray is drawn by connecting the mark with the normal using a ruler. The angle between the incident ray and the normal line is measured using a protractor. This process is continuous and is repeated with different temperatures until the water cooled down to room temperature (25°C). However, after two repeats of the investigation, it is clear that there is no observable change in the critical angle as water cools down. The marked incident rays overlap entirely. There is no need of further calculations to derive the refractive index of water and the speed of electromagnetic wave.

To confirm this conclusion, another set up is used (figure 2). The second set up observes the change in angle over a longer distance of 2 meters, so that the small changes are more noticeable. Water is boiled in a kettle and poured into the refraction dish. Similarly, the monochromatic red laser is fired at an angle of 30° to the normal to achieve refraction. A digital thermometer is used immediately after the refraction

to measure the current temperature of the water. In this particular set up, the incident ray and refraction ray were drawn using a ruler, after marking the position where the laser enters the dish. The angle of incidence and refraction is measured using a protractor. The process is repeated as water cools down. There is no observable change in the angle of refraction, which resonates with the conclusion made previously.

In this investigation the same monochromatic red laser is used for both setups.



Figure 2, refraction of laser passing through refraction dish filled with water.

3. Results

Critical angle in figure 1, angle of refraction in figure 2 were noted. To get the refractive index and wave speed, a series of calculations were carried with the variables and a variety of equations were used.

Below is used in figure 1, to derive refractive index:

$$\sin(c) = \frac{1}{n}$$

Where c in the formula above represents the critical angle.

Below is used in figure 2, to derive the refractive index:

$$n = \frac{\sin(i)}{\sin(r)}$$

Below is used in both figures, to derive the speed of electromagnetic wave speed by rearranging the formula:

$$n = c/v$$

Where the ‘c’ in the formula above represents the speed of light.

Nevertheless, data measured from the second set up is not accurate enough due to the unobservable change therefore no calculations could be made.

Table 1 : data derived from figure 2; a table showing the change in angle of refraction as temperature vary.

temperature (+/- 0.1°C)	angle of refraction (+/- 0.1°)
100	18.5
90	18.5
80	18.5
70	18.5
60	19
50	18.5
40	15 ^[1]
30	18

On the opposite hand, the secondary results found from the book: Kaye and Laby – Tables of Physical and Chemical Constants (Kaye and Laby, 1995) provided accurate data. As in table 2, there is a minor decrease in the water refractive index as the temperature increases by 5°C.

Table 2. a table showing the change in refractive index of water using light sources of different wavelengths, under temperature from 20°C-25°C. (Kaye and Laby, 1995)

Wavelength nm	Distilled water	
	$n^{20\text{ }^\circ\text{C}}$	dn/dt per $^\circ\text{C}$ (20°C-25°C)
404.66	1.342 742	-0.000 101
435.84	1.340 210	-0.000 100
486.13	1.337 123	-0.000 099
546.07	1.334 466	-0.000 098
587.56	1.333 041	-0.000 097
589.30	1.332 988	-0.000 097
632.80	1.331 745	-0.000 096
656.28	1.331 151	-0.000 096
706.52	1.330 020	-0.000 095

4. Discussion

In conclusion, there is a change in the speed of electromagnetic wave under the change in temperature. That is, as the temperature increases, the speed of electromagnetic wave also increases. There is a positive relationship between the two variables. This disproves the original hypothesis and contradicts with the photon-collision theory. Hence, when light is traveling through a medium, photons do not collide with water molecules.

Nevertheless, it should be noticed that the change in speed is insignificant and it could not be observed using both setups. There are several reasons that could explain this observation, which explore mechanism of light slowing down in a medium and the atomic arrangements when a liquid is heated.

Firstly, there is a change in speed of electromagnetic wave under the effect of temperature due to the wave-like behavior of electromagnetic waves. A theory explained by Fermilab (Fermilab, 2019) uses the wave nature to explain light’s behavior when traveling through liquid. This theory accommodates the idea that electrons in the medium are affected by the passing light. Since light is the changing of electric fields, when it passes through a medium made of atoms that are surrounded by electrons, the electrons feel a force from the oscillating electric field and starts to move. The moving electric charges further set up their own oscillating field, which is thus a wave induced by light. The induced wave combines with the light to form a new wave that travels at a slower speed through superposition. Applying the increase in temperature as a factor onto the medium, it allows electrons in an atom start to move in a more vigorous way since there is an increase in the atom’s kinetic energy. Combining the influence of temperature on electrons with the earlier theory, there is a great possibility that when temperature of a liquid is increased while light travels through it, the electrons move faster due to both effects and the wave that is induced by the light is different than a wave induced in the medium under a lower temperature. Likewise, the superposition of the new induce wave with the light generates a wave that travels at a greater speed.

Secondly, the observation could be explained by

the change in density of medium when it experiences temperature change. As temperature increases, the density of a liquid decreases due to the thermal expansion of atoms inside the liquid. The relationship between temperature and density of liquid may be represented by:

$$\rho = \frac{\rho_0}{1 + \alpha \Delta T}$$

Where ρ_0 refers to the material density at reference temperature kg/m^3

ρ refers to the material density due to temperature change in kg/m^3

α refers to the material thermal expansion coefficient in Kelvin

ΔT refers to the change of temperature in Kelvin.

Since light waves travel fastest in vacuum which has zero density and decreases in speed as it travels through a medium that has density larger than zero, the decrease in density in liquid when heated suggests an increase in the speed of light wave. Thus, density of one medium, in this case density of water, is a factor that affects the speed of light waves.

In addition, the speed change of light waves (electromagnetic wave) is insignificant as the change in critical angle is not observable in the set up due to the narrow range of temperature being investigated in figures 1, 2, and table 1.

The results and conclusions derived from the investigating water could be assumed to be applied for all liquids as they perform similar behaviors when experiencing a temperature change. However this could not be applied to any solids or gasses due to their behavioral difference to a liquid.

It should be noted that density, one of the factors deciding the speed of light waves in a medium is dependent on both temperature and pressure. Since this project had already investigated on temperature, the next step should be investigating the effect of pressure to water while light travels through to confirm the influence of changing density of water to the speed of light wave.

Theoretically, in figure 2 requires the set up to be placed between approximately 57 kilometers to gain

an observable change (1mm, the minimum distance that naked eye can observe) of the angle of refraction. This approximation could be derived using the law of refraction to find the angle of the light would bend using table 2, then use this angle to find the length that would give a 1mm horizontal distance. However this new set up is technically impossible due to the investigator's limited conditions.

In response to the refractometry industry, the conclusion of this investigation suggests that for rough calculations in the industry, the effect of temperature could be neglected. However, for fine calculations, the effect of temperature should be considered since liquids will experience slight change in refractive index under temperature change.

Acknowledgment

This investigation is completed only successfully under the assistance of the teachers of Physics: Mrs E.L.Ellis and Dr D.D.Boyce. With thanks to their suggestions and many discussions during the investigation. An additional gratitude for technician Mr D.B.Ferguson for providing laboratory apparatus for this particular project.

References

Kaye, G. W. C., & Laby, T. H. (1995). Tables of Physical and Chemical Constants and Some Mathematical Functions (Subsequent ed.). Longman Sc & Tech.

Mettler-Toledo International Inc. (2021, February 2). Refractive Index: All You Need to Know. https://www.mt.com/my/en/home/applications/Application_Browse_Laboratory_Analytics/Refractive_index/definition_and_measurement.html

Why does light slow down in water? (2019, February 20). YouTube. <https://www.youtube.com/watch?v=CUjt36SD3h8&t=0s>

Analysis of Mutation Pathogenicity and the Viability of HSP Therapy for Mutated HEX-A in Tay-Sachs Disease

Aditya Shrinivasan^{1*}

¹Evergreen Valley High School, San Jose, CA, USA

Received June 22, 2022; Revised December 13, 2022; Accepted, December 28, 2022

Abstract

This work explored three different concepts. First, docking studies were performed with various mutant HexA structures and Arimoclomol (a Heat Shock Protein inducer), with statistical analysis to establish any correlation between noted binding affinity and either mutation pathogenicity or mutation type. This was followed by the review of Arimoclomol and Heat Shock Therapy as a potential therapeutic option for Tay Sachs disease, exploring the possibility for a future clinical trial. Finally, a gene mutation pathogenicity prediction model was developed using classification with the available dataset of HEXA gene mutations to experimentally determine the pathogenicity of any HEXA mutation. The statistical analysis found no correlations between either the mutation type or mutation pathogenicity and binding affinity. This leads to the conclusion that amino acid alterations don't play a role in causing pathogenicity and benignity in a mutation and that the mutation type doesn't affect the strength of interaction between a potential treatment and the mutant protein. The mutation pathogenicity prediction model study indicated that due to the lack of sufficient features and further compounded by the low correlation between the few features, the accuracy of the resulting model was not very high. Additionally, Arimoclomol was recommended for a clinical trial with Tay Sachs Disease.

Keywords: Tay Sachs, Arimoclomol, Heat Shock Therapy, Docking, Statistical Analysis

1. Introduction

1.1 Tay Sachs Disease

Tay Sachs disease is a genetic condition caused by a mutation to the HEX-A gene preventing it from creating functional β -Hexosaminidase A (Bergeron, n.d.). β -Hexosaminidase A is a protein responsible for breaking down gangliosides, molecules that accumulate on cell surfaces in the nervous system and are broken down by the lysosome (Solovyeva, et al., 2018). Tay Sachs disease is one of several Lysosomal Storage Disorders, conditions that involve the inability to break down toxic materials such as gangliosides (Lemieux, et al., 2006). Patients with

Tay Sachs disease are unable to express β -Hexosaminidase A to at least 10% of the level of healthy humans, leading to gangliosidal buildup in the brain and spinal cord and the destruction of neurons as a part of GM2 Gangliosidosis (Suzuki, 2014).

This paper sought to explore both the effect of both mutation type and pathogenicity upon the binding affinity of a collection of mutant β -Hexosaminidase A strands with Arimoclomol, a commercially produced inducer of Heat Shock Protein— using Molecular Docking studies with Statistical Analysis (Correlation Tests). Additionally, a mutation pathogenicity prediction model was created using Machine Learning Algorithms to

* Corresponding Author
adityashrinivasan@outlook.com

Advisor: Rebecca Taylor
taylorre@esuhsd.org

predict the pathogenicity of future mutations, using the collection of available β -Hexosaminidase A mutations on the internet. In the writing of this paper, it was hypothesized that since existing Tay Sachs disease therapies do not differentiate in treatment methodology for mutation type, no correlation was expected between the type of mutation and binding affinity. Furthermore, due to the singular point of intervention present in most point mutations, no correlation was expected between pathogenicity and binding affinity.

Hypothetical results confirming the hypothesis would have enabled the consideration of Arimoclomol (or other Heat Shock Protein inducers) as a widespread potential treatment for Tay Sachs disease, due to perceived equal performance across a variety of different genotypical conditions. On the contrary, results negating the hypothesis would have called into question the methodology behind current treatment efforts, opening up possibilities for mutation-specific treatments developed independently and uniquely.

1.2 Heat Shock Protein

Heat Shock Proteins (HSPs) are a class of molecular chaperones that are responsible for regulating misfolded proteins, assisting in growth, and preventing degradation. They do this by binding with faulty proteins and assisting in their repair (Ingemann and Kirkegaard, 2014). Naturally produced by the bodies, HSPs contain disease-specific signaling processes, highlighting their ability to act as inflammation-suppressing agents exclusively (Dukay, et al., 2019). Regarding LSDs such as Tay Sachs Disease, Heat Shock Therapy is considered potentially promising due to the role of the lysosome in regulating homeostasis. With that organelle impaired, HSP Therapy and the chaperone capabilities of HSPs can maintain that homeostasis, additionally supporting deregulated proteins essential throughout the body (Miller and Fort, 2018).

This paper specifically explored the role of Arimoclomol, an HSP amplifier, in generating the Heat Shock Response throughout the body by triggering the production of specific Heat Shock

Proteins (ALS News Today, n.d.). Currently being tested with Niemann-Pick disease type C (NPC), Arimoclomol shows promise as an instigator of protein re-folding and appreciation in lysosomal function (Susman, 2021).

The study used Molecular Docking and Machine Learning. Molecular Docking is a tool that enables researchers to predict the binding positions (modes) of a particular protein-ligand combination. In this experiment, a docking software called Autodock Vina (Trott and Olson, 2010) was used to predict the binding affinity and modes of that β -Hexosaminidase A and Arimoclomol complex, revealing potential correlations between binding affinity and either mutation type or pathogenicity. The machine learning model utilized features known from existing mutation data in public repositories to predict the classification of mutations as Pathogenetic or Benign. This allowed for the development of links between specific factors and the prevalence of pathogenetic mutations leading to Tay Sachs.

2. General Methods

2.1 Molecular Docking

Molecular Docking required several components, each of which was independently prepared. At its core, docking needs a protein, a ligand, and a Grid Box. The Grid Box exists to select a region of the protein file to dock with the ligand. In this case, the protein became the mutated copies of a HexA structure found online, while the ligand was the proposed treatment- an Arimoclomol structure.

To go about the procedure, various tools were used to prepare the components for the docking. A total of 64 simulations were performed, requiring each of the 32 protein files to be properly converted. The ligand used for each was the Arimoclomol file from the PDB (CID: 9568077), while the proteins all varied from the default HexA structure (code 2GJX).

First, each of the mutated proteins was created from the base copy of the HexA protein. A list of mutations was selected from the National Institute of Health's (NIH's) Clinvar database, (Search: tay sach's and HEXA). In total, 32 mutations were selected, 18 being missense (14 Pathogenetic, 4 Benign) and 14 being Pathogenetic deletions. Using Pymol's

mutagenesis wizard, a tool to mutate downloaded structures, each of the mutations was performed. The deletions were also completed with Pymol, using the substitution feature to change each amino acid at the point of the deletion to the corresponding change. Given that there was no way to insert a stop codon, only deletions that had amino acid replacements were selected among the 14.

Several processes were used in the preparation of the 32 protein files. The downloaded .pdb files were cleaned up with standard procedures using AutoDock Tools, including deleting waters, removing non-amino acid chains, adding polar hydrogens, adding charges, and spreading the charge deficit. These processes allow for smoother docking. The ligand file, after being converted to a .pdb format, was then loaded into each docking simulation with each protein file.

The size of the Grid Box was limited, leading to the establishment of 2 docking simulations for each mutation (a total of 64). Each Gridbox covered roughly half the protein, explaining why 2 were used to test the entirety of the structure. Each simulation required the creation of a Config File, an Output File locating the contextual position of the Grid Box, and a folder containing the protein and ligand files. Utilizing AutoDock Vina, docking was executed for each mutation twice, with the Position and Binding Affinity results collected and consolidated.

Utilizing SPSS statistics, the collected data was sorted and tested for correlations between two sets of variables: Mutation Type v Binding Affinity and

Pathogenicity v Binding Affinity. Again utilizing SPSS, a scatter plot was created to visualize these results.

2.2 Mutation Model

The purpose of the mutation model was to use established Tay Sachs mutation pathogenicity, type, and location data to predict the pathogenicity of new mutations or mutations of unknown pathogenicity. In order to do so, the list provided on ClinVar was downloaded, with feature consolidation including the establishment of a unique ID for model-making purposes and the addition of Mutation Type and Allele Change Classification for Single Nucleotide Polymorphisms from NIH information.

The dataset was then split into a test and training set, with basic statistical models (Log Reg, Nearest Neighbour, Support Vector Machines (SVMs), Kernel SVMs, Bayes, Decision Trees, and Random Forest) run against the data.

Several trials were run using different alterations to the dataset. Experimentally, predict2snp, a mutation pathogenicity prediction tool, was used to classify mutations that on ClinVar were listed with unknown pathogenicity in trial 5. Trials 2, 3, and 4, instead, classified all mutations of unknown pathogenicity as benign, while trial 2 only utilized missense mutations, eliminating the others from the dataset. Trial 1 was the only trial utilizing 3 distinct classifications (P, B, and Unknown). Trial 5 was referred to as the “Golden Trial” in the results.

3. Results

3.1 Docking Studies + Statistical Analysis

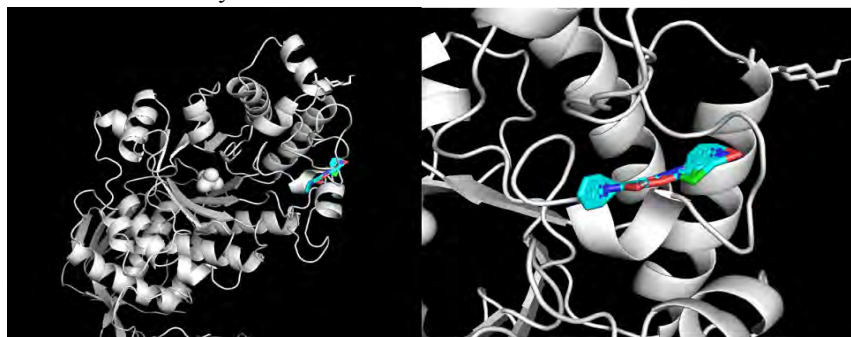


Fig. 1: Most optimal of 9 proposed binding positions (zoomed in on the right) for Arimoclomol with the mutated Hex-A (deletion of the 47th Amino Acid)

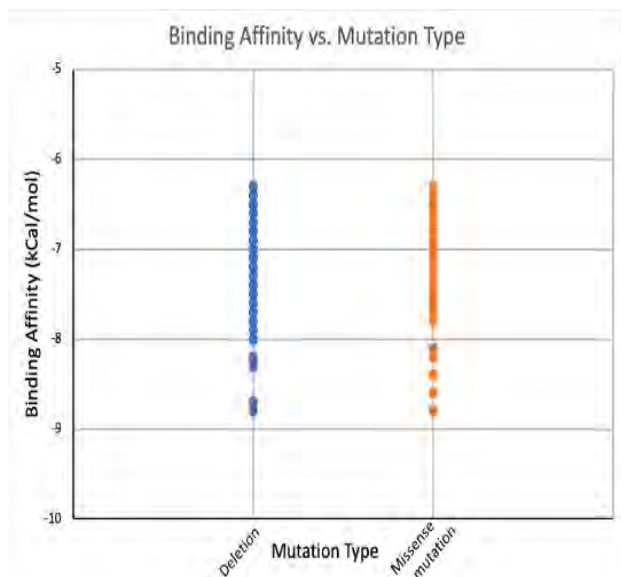


Fig. 2: Binding Affinity plotted vs Mutation type

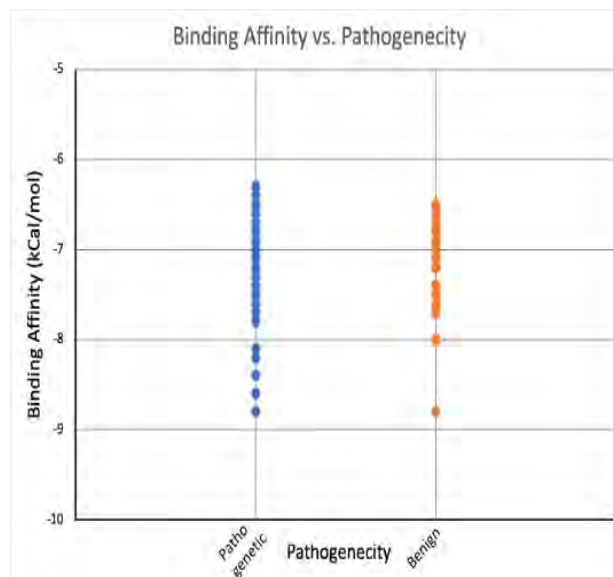


Fig. 3: Binding Affinity plotted vs Pathogenicity

Table. 1: Correlation between Binding Affinity, Mutation type, Pathogenicity

Correlation between Binding Affinity and Mutation Type			
		Mutation Type	Affinity
Mutation Type	Pearson Correlation	1	0.021
	Sig (2-tailed)		0.634
	N	504	504
Affinity	Pearson Correlation	0.021	1
	Sig (2-tailed)	0.634	
	N	504	504
Correlation between Binding Affinity and Pathogenicity			
		Mutation Type	Affinity
Pathogenicity	Pearson Correlation	1	-0.074
	Sig (2-tailed)		0.183
	N	324	324
Affinity	Pearson Correlation	-0.074	1
	Sig (2-tailed)	0.183	
	N	324	324

3.2 Mutation Pathogenicity Prediction Model

Table 2. Performance of different Machine Learning Algorithms in predicting Pathogenicity of mutations

Parameter		Golden Trial	Trial 4	Trial 3	Trial 2	Trial 1
Dataset size		441 entries, 5 features	439 entries, 5 features	441 entries, 5 features	384 entries, 4 features	384 entries, 4 features
Characteristics of dataset		All mutations of uncertain pathogenicity were classified using predicti2snp prediction tool, included insertion and inversion mutation types.	All mutations of unknown uncertain were classified as benign, no inversions or insertion mutation types were included in the dataset.	All mutations of unknown uncertain were classified as benign, indel, duplication, and deletion mutation types included in the dataset.	All mutations of unknown uncertain were classified as benign, only missense mutations were in the dataset. Mutation Type feature not included.	3 categories of diagnosis: Pathogenetic, Benign, or Uncertain. Only missense mutations were in the dataset, without Mutation Type as a feature.
Algorithm Accuracy	Logistic Regression	63%	65%	65%	75%	43%
	Nearest Neighbor	56%	68%	64%	73%	46%
	Support Vector Machines	56%	64%	65%	75%	44%
	Kernel SVM	63%	61%	68%	73%	46%
	Naïve Bayes	45%	65%	58%	75%	41%
	Decision Tree Algorithm	59%	63%	62%	73%	49%
	Random Forest Classification	56%	63%	64%	71%	47%

4. Discussion

4.1 Docking Studies + Statistical Analysis

The goal of the statistical analysis was to reveal correlations- represented by the Pearson Correlation Coefficient, a number from (-1) to (1) constituting the strength and direction of the relationship between two variables. Paired with the significance- another number that reveals the probability of observing a specific correlation due to chance alone (lower significance is preferable), the resulting numbers can be used to create observable conclusions regarding the presence of accurate and impactful correlations.

With the observed results, the most pressing takeaway would be the noted lack of correlation between either the mutation type and binding affinity or the mutation pathogenicity and the binding affinity. These discoveries line up with the

hypothesis. Figures 1, 2, and 3 display the arimoclomol-mutated Hex-A interaction and the graph of the binding affinity values first versus pathogenicity of missense mutations and second versus distinct pathogenetic mutation types.

In the first case, mutations were sorted into only missense mutations and deletions, given that those were the two most commonly occurring types of mutations, and performing insertion mutations with Pymol was difficult. The second case sorts missense mutations of benign and malignant pathogenicity against the tested binding affinity, only testing missense mutations for internal control. Table 1 (IBM, n.d.) displays results from the Pearson's Correlation Coefficient Test, comparing Binding Affinity with both Mutation Type and Pathogenicity. The 2 Tailed Significance, being over the accepted 0.001, implicates a lack of correlation between either test, suggesting that neither mutation type nor

mutation pathogenicity plays a key role in the binding affinity calculation. As visualized in Figure 2, the distribution of binding affinities across both mutation types indicates that mutations vary in strength of attraction with the treatment. Given that binding affinity correlates with the strength in response to the treatment, these results indicate that neither mutation type tested has a stronger response to the Arimoclomol, and that response to the Arimoclomol varies across patients with the same mutation type.

In the second case (Figure 3), pathogenicity was tracked for binding affinity, holding the type of mutation as constant: all missense mutations. The lack of correlation observed emphasizes the finding that amino acid changes that constitute a pathogenetic variant are no less likely to have a stronger or weaker interaction than amino acids that constitute a benign mutation. It can thus be concluded that in addition to pathogenicity, amino acid changes are not correlated with the binding affinity of the Arimoclomol.

4.2 Mutation Model

Table 2 displays the trials and their corresponding results. The mutation pathogenicity predictive model suffers from a lack of features and entries. Working with an incredibly small dataset- under 500 total entries- and without many or uniquely categorized features (just position data, pathogenicity, mutation type, and factual classifications), any results must be scrutinized with the understanding of the nature of classification problems and the small dataset (Goel, 2018).

This means that although certain trials gave a higher accuracy rating, the manipulation of variables to get that accuracy rating affects the reliability of the methodology of the model. In theory, by reducing the number of features, the classification becomes increasingly binary, reliant upon correlations between a few features. If there exists a strong correlation between those features, then the accuracy rating would be high. However, if there doesn't exist a strong correlation, then the accuracy rating is much lower (ProjectPro, 2022). In either case, the model doesn't consider the holistic picture, focusing on only

a few factors that individually may or may not have a high correlation with the pathogenicity of the mutation. Ideally, the model should consider a wide range of factors, analyzing all possible variables affecting mutation pathogenicity.

Due to the limited data, the model lacks the features required to make those distinctions, instead of searching for a correlation between a few individual features. Since none of the few noted features had a strong correlation, the accuracy was quite low. Still, the developed methodology, despite the lower predictive score, is indicative of the process carried out with larger data sets, outlining the processes that should be followed for the long-term development of the tool as more mutations are classified into databases and more features are categorized by researchers over time.

In the trials provided, the most complete trial would be the golden trial, which incorporated the usage of predict2snp (Bendl, et al., 2014) to classify mutations of uncertain pathogenicity as either benign or pathogenetic. Within this trial, either the Logistic Regression or Kernel Support Vector Machine yields the highest percentage- at 63% accuracy. However, the highest pure accuracy would be in the 2nd trial, with mutations of uncertain pathogenicity classified as benign, non-missense mutations excluded completely, and mutation type not yet added as a feature. The 75% reflected with the Logistic Regression, Support Vector Machine, and Naïve Bayes method reflects upon the increasingly binary nature of the classification in that problem, as there were fewer features for the model to consider- illustrating the point being made about small datasets and a limited feature set per mutation.

The methodology in the golden trial, when applied to a larger dataset with more features, would reflect the sought procedure for the development of a mutation prediction model. The conclusions to draw from these results are based on the need for a more comprehensive categorization of data. With the inclusion of a few physical features, such as the dimensions and size of the noted cherry-red spot characteristic of Tay Sachs, the model would have more factors to analyze and would become an effective tool in the immediate classification of a HEXA gene mutation as either benign or

pathogenetic. Still, at its current effectiveness, the model does not successfully predict mutation pathogenicity.

4.3 Heat Shock Proteins and HSP Therapy

Heat Shock Proteins have emerged as a potential target for therapeutic intervention with several Lysosomal Storage Disorders, such as Niemann-Pick Type C (NPC). Specifically, Arimoclomol, a promoter of HSP Expression, has been explored with NPC, with promising results. HSPs are desirable for their induced responses- among which include the Heat Shock Response, a nerve response that is triggered by the body to combat various stressors such as protein misfolding (ALS News Today, n.d.).

The therapeutic capabilities of the HSR response can protect against degenerative disorders, such as Lysosomal Storage Diseases. LSDs involve the buildup of toxic substances due to a lack of specific enzymes (Orphazyme, 2021). Aggregate conditions caused by LSDs include protein misfolding- such being the case in both NPC and Tay Sachs. The capabilities of various HSPs and their capacity for controlling protein misfolding in addition to the HSR's ability to regulate cell death, autophagy, membrane permeabilization, and aid in controlling various other cellular processes make them an attractive choice for LSDs.

In mice, Recombinant HSPs have been shown to mitigate the effects of NPC, providing conclusive evidence that HSPs can aid with disorders affecting the lysosome (Susman, 2021). However, in humans, Recombinant HSPs struggle to cross the Blood-Brain Barrier, meaning that the next best solution is Arimoclomol- an HSP regulation enhancer. In mice, Arimoclomol causes co-localization of HSPs in the cerebellum with their target enzyme, allowing the "fixing" process to initiate protein re-folding. In humans, Arimoclomol was shown to mitigate disease progression in patients transferred from the placebo arm of the trial to the Arimoclomol treatment. In general, the associated NPCCSS score between the two treatments favored Arimoclomol as the more effective option, with a "65% relative reduction in annual disease progression." (Mengel, et al., 2021).

Researchers speculate that Arimoclomol in

tandem with a drug that slows disease progression (such as miglustat- an inhibitor of glucosylceramide synthase- in NPC) could be the best approach for targeting NPC in humans (Susman, 2021).

Tay Sachs Disease shared several similarities with NPC, from the fact that both are autosomal recessive Lysosomal Storage Diseases to the misfolding of proteins that cause symptoms in both (National Organization of Rare Disorders, n.d.). The promise of HSPs as molecular chaperones inducing refolding efforts, autophagy, and cell permeabilization extends to Tay Sachs, where Arimoclomol would be an intriguing option to test with a clinical trial.

In the absence of a reliable gene therapy and given the doubts over the efficacy and deliverability of chaperones such as Pyrimethamine (Parenti, et al., 2015), a chaperone inducer such as Arimoclomol could help deliver the same desired effects as Recombinant HSPs without the issues in deliverability.

5. Conclusion

In the docking study and ensuing statistical analysis, neither mutation type nor pathogenicity correlated with binding affinity, meaning that the effectiveness of Arimoclomol as a treatment for Tay Sachs disease would be expected to remain the same regardless of disease type. Additionally, the study with pathogenicity confirmed that the factors affecting binding affinity do not include the amino acid changes, as unique affinities were not observed for benign or pathogenetic variants.

With regards to the mutation pathogenicity prediction ML model, the methodology has been set for future feature addition and the addition of further entries over time. An interesting exploration would be the derivation of physical features explaining physically noted traits (for example, the size of the famous red spot in the back of the eye indicative of Tay Sachs). The accuracy is understandably low, in contrast to the Hypothesis, given that all public datasets lacked data beyond the 5 features derived, of which one was the unique ID. The size of the dataset- under 500 entries for each trial, also limits the size of the test and training sets. The potential with such a model, to immediately categorize Tay Sachs

mutations and provide additional time for advancement treatment, is immense, and the classification of additional features in the mutation diagnosis stage will only help it improve.

In conclusion, Arimoclomol remains an exciting option to be tested for a clinical trial for Tay Sachs Disease. The existence of testable animal models means that the search for a cure for Tay Sachs is accelerating, and the end product is closer than ever. Given Arimoclomol's success with Niemann Pick Type C, another Lysosomal Storage Disease, it is an ideal candidate for a clinical trial with Tay Sachs.

Acknowledgment

I would like to thank Mrs. Rebecca Taylor of Evergreen Valley High School for her help and mentorship during this research work.

References

ALS News Today. (n.d.). *Arimoclomol*.
<https://alsnewstoday.com/arimoclomol-orph-001/>

Bendl, J., et al. (2014). PredictSNP: robust and accurate consensus classifier for prediction of disease-related mutations. *PLoS computational biology*, 10(1), e1003440.
<https://doi.org/10.1371/journal.pcbi.1003440>

Bergeron, S. (n.d.). *Tay-Sachs*.
http://www-personal.umd.umich.edu/~jcthom/JCT_HOMAS/1997_Case_Studies/S_Bergeron.html

Dukay, B., Csoboz, B., & Tóth, M. E. (2019). Heat-Shock Proteins in Neuroinflammation. *Frontiers in pharmacology*, 10, 920.
<https://doi.org/10.3389/fphar.2019.00920>

Goel, V. (2018, September 29). *Building a Simple Machine Learning Model on Breast Cancer Data*. Towards Data Science.
<https://towardsdatascience.com/building-a-simple-machine-learning-model-on-breast-cancer-data-eca4b3b99fa3>

IBM Corp. (n.d.). *IBM SPSS Software*. IBM Corp.
https://www.ibm.com/analytics/spss-statistics-software?utm_content=SRCWW&p1=Search&p4=43700068092265694&p5=p&gclid=CjwKCAjwCjwKCAjwCjwKCAjwCjwKCAjwAT1fY7xIpmbUD0Wqo98dYleTtElf5NJWjCwfOL

[hj7Z6rGQc_Vhl_93CMzpBoCZs0QAvD_BwE&gclid=CjwKCAjwCjwKCAjwAT1fY7xIpmbUD0Wqo98dYleTtElf5NJWjCwfOL](https://doi.org/10.1194/jlr.R048090)

Ingemann, L. & Kirkegaard, T. (2014). Lysosomal storage diseases and the heat shock response: convergences and therapeutic opportunities. *The Journal of Lipid Research*, 55(11), 2198–2210.
<https://doi.org/10.1194/jlr.R048090>

Lemieux, M. J., et al. (2006). Crystallographic structure of human beta-hexosaminidase A: interpretation of Tay-Sachs mutations and loss of GM2 ganglioside hydrolysis. *Journal of molecular biology*, 359(4), 913–929.
<https://doi.org/10.1016/j.jmb.2006.04.004>

Mengel, E. et al. (2021). Efficacy and safety of arimoclomol in Niemann-Pick disease type C: Results from a double-blind, randomised, placebo-controlled, multinational phase 2/3 trial of a novel treatment. *Journal of inherited metabolic disease*, 44(6), 1463–1480.
<https://doi.org/10.1002/jimd.12428>

Miller, D. J., & Fort, P. E. (2018). Heat Shock Proteins Regulatory Role in Neurodevelopment. *Frontiers in neuroscience*, 12, 821.
<https://doi.org/10.3389/fnins.2018.00821>

National Organization of Rare Disorders. (n.d.). *Rare Disease Database – Tay Sachs Disease*.
<https://rarediseases.org/rare-diseases/tay-sachs-disease/>

Orphazyme. (2021, March 3). *About Heat Shock Proteins (HSPs)*.
<https://www.orphazyme.com/our-science/>

Parenti, G., Andria, G., & Valenzano, K. J. (2015). Pharmacological Chaperone Therapy: Preclinical Development, Clinical Translation, and Prospects for the Treatment of Lysosomal Storage Disorders. *Molecular therapy : the journal of the American Society of Gene Therapy*, 23(7), 1138–1148.
<https://doi.org/10.1038/mt.2015.62>

ProjectPro. (2022, May 24). *Classification vs. Regression Algorithms in Machine Learning*.
<https://www.projectpro.io/article/classification-vs-regression-in-machine-learning/545>

Solovyeva, V. V., et al. (2018). New Approaches to Tay-Sachs Disease Therapy. *Frontiers in physiology*, 9, 1663. <https://doi.org/10.3389/fphys.2018.01663>

Susman, E. (2021, October 6). Arimoclomol
*Appears to Slow Progression of Niemann-Pick Type
C Disease. Neurology Today.*
<https://journals.lww.com/neurotodayonline/blog/neurologytodayconferencereporterscnsannualmeeting/pages/post.aspx?PostID=48#:~:text=October%206%2C%202021&text=%E2%80%8BTreatment%20with%20the%20investigative,Child%20Neurology%20Society%20annual%20meeting>

Suzuki, Y. (2014). Emerging novel concept of

chaperone therapies for protein misfolding diseases.
*Proceedings of the Japan Academy. Series B,
Physical and biological sciences*, 90(5), 145–162.
<https://doi.org/10.2183/pjab.90.145>

Trott, O., & Olson, A.J. (2010). AutoDock Vina:
improving the speed and accuracy of docking with a
new scoring function, efficient optimization, and
multithreading. *Journal of computational chemistry*,
31(2), 455–461. <https://doi.org/10.1002/jcc.21334>

Breakthroughs in Honey Bee Health, Continuous-Release Mist Diffusion of Thymol-Based Essential Oils: Part II - The Field Study

Kaitlyn N. Culbert¹*

¹Toms River High School North, Toms River, NJ USA

Received July 24, 2022; Revised December 20, 2022; Accepted, January 3, 2023

Abstract

Honey bees (*Apis mellifera*) are pivotal pollinators in agricultural and natural ecosystems; however, since the winter of 2006-2007, honey bee colonies have been disappearing. The greatest single contributor to the decline of honey bee health is the *Varroa* mite. Synthetic chemicals are used to control *Varroa*, but the mites are developing resistance. Essential oils (EO) may be a viable alternative. EO are cheaper, environmentally-friendly, pose fewer health risks to bees and consumers, and *Varroa* have not developed resistance to the oils. Currently, all commercially available thymol-centered systems are gel-based and work by direct contact with the mite. These systems are also highly dependent on temperature and humidity for effectiveness. Following the laboratory investigation (Part I), this field study (Part II) examined the use of thymol-based EO for miticide efficacy as well as the use of mist diffusers to eliminate any dependence of the essential oils on temperature and humidity conditions. Miticide efficacy was recorded as: thyme>oregano>rosemary>control (vegetable glycerin). Across all tested EO, the highest miticide activity occurred during the first two weeks of treatment. A brief cost analysis demonstrated using mist diffusers was more cost-effective than commercially available thymol-based systems (US\$3.20 versus US\$15-\$18 per application). Continuous-release mist diffusion permits the disbursement of EO throughout the entire hive and effectively provided early elimination of mites as they emerged from the brood cell, while remaining safe for honey bees.

Keywords: Varroa Destructor; Apis Mellifera; Mites; Honey Bees; Essential Oils; Mist Diffusion

1. Introduction

Honey bees, *Apis mellifera*, are crucial pollinators for agriculture, responsible for over 80% of all cultivated crops (Randall, 2020). In fact, bee pollination accounts for approximately US\$15 billion in added crop value (USDA, 2021). Additionally, bees also produce honey, pollen, royal jelly, beeswax, propolis, and venom for nutritional and medicinal uses for an additional US\$300 million annually (Calovi, 2021). In the United States, there are approximately 2.5 million commercially farmed honey bee hives and around 500,000 colonies kept by

hobbyists and semi-professional bee keepers (Penn State Extension, 2013). Unfortunately, these numbers are declining at a rapid rate. Losses are attributed to Colony Collapse Disorder which occurs when there is a sudden loss of a colony's worker bee population yet the queen, brood and a relatively abundant amount of honey and pollen reserves remain. Various reports have suggested losses between 30% to 50% of winter bee colonies in the US, (EPA, 2021); its lowest point in the past 50 years.

Honey bee colonies do not remain dormant during the winter and remain active to maintain the hive temperature between 24-34 degrees Celsius

* Corresponding Author
culbertkaitlyn@gmail.com

Advisor: Christine Girtain
cgirtain@trschoools.com

(75.2-93.2 degrees Fahrenheit) by forming a thermoregulating cluster (Calvoli, 2021) which enables them to survive long periods of cold temperature. The colony relies on its existing honey stores and stops foraging for nectar and pollen and brood rearing ceases. While summer worker bees live only for a few weeks, winter bees live for several months. A critical component to increasing the lifespan of winter bees is the control of *Varroa destructor* mites. The greatest single contributor to the decline of bee health is the *Varroa* mite (Rosencranz, 2010). A single *Varroa* mite can shorten the lifespan of a bee by one-third, and two mites can shorten it by one-half (Bryant, 2006). *Varroa* weakens and ultimately kills colonies by out-reproducing their host.

Varroa is an ectoparasitic mite that feeds on the fat bodies of developing honey bee larvae and adult bees and aggressively reproduces within an infected bee colony. Recent research by Ramsey et al. brings to light the *Varroa* mite's focus on the fat body tissue, and not the hemolymph as previously believed, rendering the honey bee susceptible to harm from pesticides (Ramsey, 2019). Furthermore, parasitism by *Varroa* decreases the body weight and water content of young emerging bees (Noel, 2020). The lowered weight of the future adult bee increases with the number of mites. Specifically, by decreasing the size of drones, *Varroa* induce a deficit in sperm production (Noel, 2020). *Varroa* also alters flying, homing and orientation abilities in foragers thereby limiting the efficiency of honey bees collect resources needed for colony development.

To control the spread of *Varroa*, bee keepers initially used acaricides, pyrethroids, and organophosphates pesticides (Bahreni, 2020). The frequent use of these synthetic miticides to control *Varroa* infestations has resulted in the development of resistance to many of the chemical components of these miticides (Bahreni, 2020, Traynor, 2016).

Essential oils may be an alternative to chemical pesticides. They are cheaper, environmental-friendly, and pose fewer risks to the health of bees and consumers. Most importantly, *Varroa* have not developed resistance to essential oils for honey bee mite control (Ghasemi, 2011; Damiani, 2009).

Currently, numerous essential oil compounds have

been evaluated for miticidal activity. One of the proven successful essential oils is thymol. It works by disorienting the mite and blocking its pores (Tennessee's Honey Bees, 2021). Thymol is the only compound of essential oils widely used in beekeeping with 70%-90% efficacy against *Varroa* (Garrido, 2018) Thymol also has the added advantage of being active against fungus (chalkbrood) and some efficacy in tracheal mites (Davis, 2003). The most widely used and popular beekeeping products with thymol as a main ingredient are Apiguard®, ApiLifeVar® and Thymovar® (Garrido, 2018). None of these currently available systems utilizing thymol reach reproducing *Varroa* mites in the brood cell (Garrido, 2018). All of these delivery systems only kill the mites on the adult bees (Garrido, 2018).

Although a wide variety of essential oils (over 150) have been tested as potential miticides; very few of them have proven successful when tested in field trials (Sabahi, 2017). Unfortunately, a key problem, evidenced in field trials to evaluate the miticide activity of essential oils, is that results are not consistent; there is tremendous variability between studies due to local temperatures and the relative humidity affecting EO efficacy (Sabahi, 2017). All commercially available thymol-based miticide systems are gel-based and the effectiveness of the thymol is highly dependent upon the ambient temperature and relative humidity. Temperature and humidity affect the rate of essential oil evaporation (Sabahi, 2017). Furthermore, because, these systems are gel-based, they are only effective when there is direct contact with the mite (Garrido, 2018). Following the laboratory investigation (Part I), this field study (Part II) examines the use of thymol-based EO for miticide efficacy as well as the use of mist diffusers to eliminate any dependence of the essential oils on temperature and humidity conditions for EO efficacy.

2. Materials and Methods

2.1. Apiary and Colonies

Following a laboratory investigation demonstrating the potential utility of battery-operated mist diffusion of thymol-based essential oils, a field

study was conducted in *A. mellifera* colonies with bee hives naturally infested by *Varroa destructor* mites located at Cherry Hill, New Jersey. The level of *Varroa* mite infestation of colonies, headed by Italian queens and housed in Langstroth hives, was determined by measuring the rate of mite drop before starting the field study. The hives were equipped with screened bottom boards and 3mm mesh hardware cloth and sticky paper coated with vegetable shortening (Crisco®, B&G Foods, Parsippany, NJ, USA) to capture falling mites. Four colonies were selected for having similar mite infestation rates. Each colony was assigned to one of four treatment groups. Trials commenced October 2, 2021 for a 4-week period and concluded October 30, 2021.

2.2 Essential Oils and Continuous-Release Battery-Operated Mist Diffusers

For this experiment, three USDA, certified organic, premium food-grade (highest grade) essential oils (Zongle Therapeutics, Norcross, GA, USA) consisting of thyme (*Thymus linearis*), oregano (*Lippia berlandieri*), and rosemary (*Rosmarinus officinalis*) were evaluated along with a control consisting of organic, premium food-grade, vegetable glycerin (Plant Guru, Plainfield, NJ, USA).

To provide continuous-release, two alternating battery-powered mist diffusers (AirWick® Essential Mist Essential Oil Diffuser, Reckitt Benckiser, Slough, England) were used in each hive to provide 24/7 release because each mist diffuser automatically shuts down after eight hours of continuous use. Battery-powered mist diffusers were selected due to the lack of electrical power available at the hive location. The AirWick Mobile App was utilized to ensure that all diffusers were operating properly to provide 24/7 continuous-release. Hive monitors (Broodminder T2SM Internal Hive Monitors, Stoughton, Wisconsin, USA) were utilized to monitor temperature and relative humidity in the hive. The mist diffusers and hive monitor set-up in the hive super are depicted in Figure 1.

2.3 Mite Mortality/Bee Mortality

To assess treatment efficacy and rate of mite fall,

sticky papers were placed underneath the screen of the hives' bottom boards to capture falling mites as described in previous field studies (Asha, 2015; Sabahi, 2017). Furthermore, honeybees are known to remove fallen or diseased nestmates. Removing corpses protects against infection. As depicted in Figure 2, the papers were collected twice a week and fallen mites on them counted and divided by 3 or 4 (depending on day of week) to obtain an average rate of mites dropped per day. For example, a mite collection and count conducted on Wednesday, would be divided by 4 to obtain a daily average mite fall for Sunday, Monday, Tuesday, Wednesday whereas a mite collection and count conducted on Saturday would be divided by 3 for an average daily average mite fall for Thursday, Friday, Saturday. After 4 weeks of trials, two plastic strips containing amitraz (Apivar, Ve'to-Pharma, Villebon-sur-Yvette, France) were placed in each hive as finisher treatments.



Figure 1. Essential Oil/Mist Diffusers Set-Up in the Hive Super

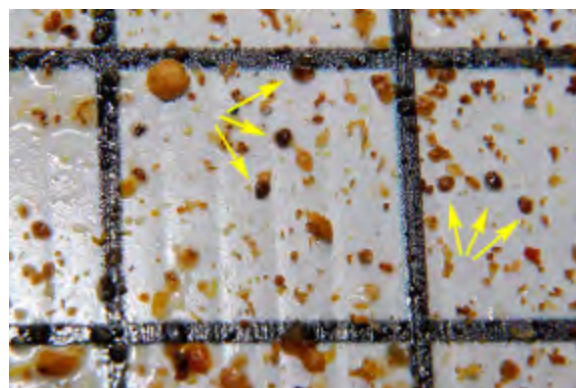


Figure 2: Fallen Mites on Sticky Paper

2.4 Treatment Efficacy

The efficacy of each treatment was determined using the following equation:

$$\text{Efficacy (\%)} = \frac{\Sigma[N1/(N1 + N2)] \times 100$$

where, N1 is the number of mites that fell during the 4 weeks of test treatments, and N2 is the number of mites that fell during the final treatment with amitraz.

Early *Varroa* mite control in the fall season is important to improve the health of colonies and to increase their likelihood of winter survival. Therefore, the proportion of mites that were killed by the treatments each week, during the treatment period, was calculated. The Relative Mite Weekly Fall Rates (MFWRs) were calculated for each treatment using the following equation:

$$\text{MFWR} = [(a + b)] / c \times 100$$

“a” is the number of mites that fell during the first observation of a week, b, is the number of mites that fell during the second observation of the same week, and c, is the total number of mites that fell during the 4 weeks of treatment.

To evaluate the effect of compounds on bee mortality, Todd dead bee entrance traps (Stoner 1979) were installed at the entrance of hives as depicted in Figure 3. Dead bees were counted twice a week during the trial’s 4-week period.



Figure 3: Todd Dead Bee Entrance Trap

3. Results and Discussion

3.1 Miticide Efficacy

As shown in Figure 4, the efficacy rates for *Varroa* mite control among the essential oil treatments were significantly higher than the

vegetable glycerin control. Treatment 2 (thyme essential oil) had the highest percent efficacy rate (97.4% ± 0.68), which was followed closely by Treatment 1 (oregano essential oil) with a comparable percent efficacy rate (93.8% ± 0.72) and more distantly by Treatment 3 (rosemary essential oil with an efficacy rate of 78.2% ± .89). Conversely, the lowest percent efficacy rate was observed for vegetable glycerin control (42.8% ± .81). Statistical significance was determined using regression analysis. Statistical significance is the probability of an observation not being caused by a sampling error. When the p-value is equal to or less than 0.05 the statistical result is believed to be accurate. All p-values ≤ 0.05 and were found to be statistically significant [*thyme* (p=.00000742); *oregano* (p=.0000593); *rosemary* (p=.000487)].

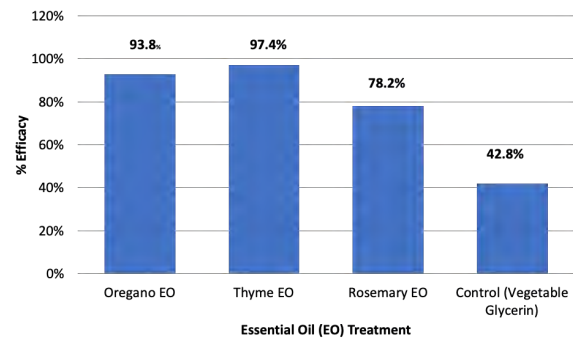


Figure 4: Mean Varroa Mite Control % Efficacy of Tested Essential Oil Treatments in Hives

3.2 Mite Fall Weekly Rates (MFWR)

As depicted in Figure 5, the mite fall weekly rates (MFWR) was dramatically higher for all essential oil treatments in the first two weeks of treatments in comparison to the vegetable glycerin control. The control had greater mite fall percentage in the last two weeks whereas thyme exhibited the highest proportion of mites killed during weeks 1 and 2 (82.6%) followed by oregano (73.2%) and rosemary (61.8%) in comparison to the control (45.7%). In this scenario, where there is a continuous release of essential oils at a sufficient dose over an extended period, it appears that the mites are killed as they emerged from brood cells. As such, a large proportion of the mites were killed within the first 2

weeks of treatment (particularly during the second week of treatment) due to miticide efficacy of the essential oils. A significant difference should be noted that in the last week of treatment across all the essential oils. During the last week of treatment, the percent of mite fall was between 4.1% to 12.3% for all EOs in comparison to the percent of mite fall of the control at 26.2%. Miticide efficacy of thymol-based essential oils was greatest during the first two weeks of treatment in contrast to the vegetable glycerin control which exhibited greater percentage of falls during the last two weeks of treatment. Statistical significance was determined using regression analysis. Statistical significance is the probability of an observation not being caused by a sampling error. When the p-value is equal to or less than 0.05 the statistical result is believed to be accurate. All p-values ≤ 0.05 and were found to be statistically significant [thyme ($p=.00048$), oregano ($p=.00929$); and rosemary ($p=.00658$)]. $n=1,281$.

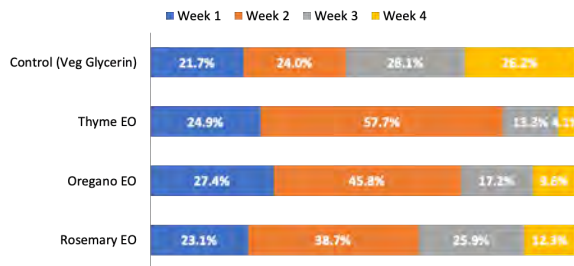


Figure 5: Mite Fall Weekly Rates of Tested Essential Oil Treatments in Hives

3.3 Worker Bee Mortality

As illustrated in Figure 6, the average per day worker bee mortality was similar across all tested essential oils and the vegetable glycerin control indicating the tested essential oils were safe for honey bees. The greatest daily mortality was found in thyme (29.3), followed by oregano (28.8), and rosemary (27.5). All of which was comparable to the vegetable glycerin control (26.8). Statistical significance was determined using regression analysis. Statistical significance was determined using regression analysis. Statistical significance is the probability of an observation not being caused by a sampling error. When the p-value is equal to or less

than 0.05 the statistical result is believed to be accurate. All p-values > 0.05 and not statistical significant [thyme ($p=.876$); oregano ($p=.594$); rosemary ($p=.452$)]. $n=836$.

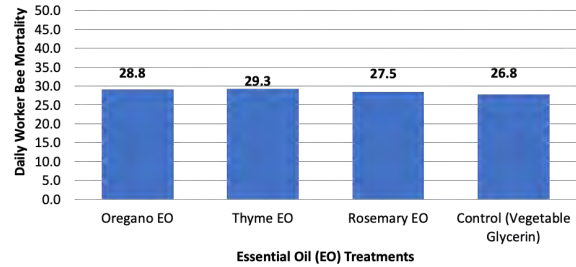


Figure 6: Per Day Worker Bee Mortality of Tested Essential Oil Treatments in Hives

4. Discussion

This investigation of various essential oils demonstrated the effectiveness of select thymol-based essential oils for the control of *Varroa* mites during the fall in the hive. All tested essential oils were shown to be effective miticides with no effect on worker bee mortality. A critical advantage of using essential oils in hives is that essential oils are generally recognized as safe and environmentally-friendly and the mites have not developed resistance to the oils. All the essential oils utilized for this field study were premium food-grade quality. Worker bee mortality was found to be the same across all essential oils and comparable in safety to the vegetable glycerin control.

This study also demonstrated the clinical utility of using a battery-operated continuous-release delivery system. Our results demonstrate that thymol-based essential oils delivered with battery-operated mist diffusers can achieve a high level of *Varroa* mite control. The most effective miticide control came from thyme (97.4% efficacy), followed by oregano (93.8% efficacy) and rosemary (78.2% efficacy) in comparison to the vegetable glycerin control (42.8%). The miticidal activity of thymol-based essential oils is attributed to terpenes like carvacrol (Sabahi, 2017). In our study, the tested oils contained between 60-69% carvacrol; however, the concentration of this component may vary from source to source and can potentially affect the essential oil's varroacidal

efficacy.

The use of a battery-operated mist diffusion system provided a continuous-release mist diffusion of thymol-based essential oils throughout the hive. In contrast, currently available thymol-based gel systems are limited by their need to be in direct contact with the mite as well as a “goldilocks” range of temperature and humidity for efficacy. The proportion of mites killed at different treatment periods can be attributed to the mist diffusion thymol-based essential oil technology. For thyme, 82.6% of the total number of mites that died during the 4-week trial period, fell during the first 2 weeks. For oregano, 73.2% and rosemary, 61.8%. However, only 45.7% of the mites fell during this time period in the vegetable glycerin control. In this scenario, where there is a continuous release of essential oils at a sufficient dose over an extended period, the mites are killed as they emerge from brood cells. As such, a large proportion of the mites were killed within the first 2 weeks of treatment (particularly during the second week of treatment) due to miticide efficacy of the essential oils. Only between 4.1%-12.3% of mites had fallen during the last week of treatment indicating the majority of mites were killed as they emerged from the brood cell. The early elimination of mites is critical as it results in longer bee lifespan and higher colony survival after winter (van Dooremalen, 2012)

Lastly, a cost-effectiveness analysis found the average cost per application of the thymol-based essential oil continuous-release mist diffuser technology is approximately US\$3.20 per hive per application (assuming the use of 2 mist diffusers) in comparison to currently available thymol-based commercial products at US\$15-US\$18 per hive per application.

5. Conclusion

In light of the current crisis surrounding honey bees and the *Varroa* mite, new solutions are desperately needed. The goal of this field study was to demonstrate that battery-operated, continuous-release mist diffusion of thymol-based essential oils can serve as a cost-effective miticide while also being safe for honey bees. Thymol-based essential oils

delivered with battery-operated mist diffusers achieved a high level of *Varroa* mite control (thyme > oregano > rosemary > vegetable glycerin control) particularly in the first two weeks of treatment allowing for a longer bee lifespan and higher colony survival after winter. Continuous-release mist diffusion of thymol-based essential oils may effectively, safely, and cost-effectively be incorporated as part of a natural miticide control plan to enhance the chances of honey bee colony survival.

Acknowledgements

I would like to acknowledge the following individuals who provided advice, guidance and access to their apiary throughout my research: Professor Michael Haberland (Rutgers), Dr. Ron Hutchison (Stockton), Dr. Jack Cameron (University of Florida), Randy Oliver (commercial beekeeper), Geff Vitale (CJBA), Cliff Moore (CJBA), Peter Van Mater (CJBA), my beekeeping mentor Angela Juffey (CJBA), and Mann Lake Ltd for their donation of beehives and equipment.

References

- Asha, P., et al. (2015). Comparative evaluation of sticky paper and hive debris as sampling methods for population assessment of *Varroa destructor* in *Apis mellifera* colonies. *Indian Journal of Agricultural Research*. 49. 10.18805/ijare.v49i6.6676.
- Bahreini, R., et al. (2020). Evaluation of potential miticide toxicity to *Varroa destructor* and honey bees, *Apis mellifera*, under laboratory conditions. *Sci Rep* 10, 21529.
- Beekeeping - Honey Bees. (2013, June). Penn State Extension. January 22, 2021, from <https://extension.psu.edu/beekeeping-honey-bees>
- Bryant, D. (2006, February 16). USDA Bee Lab Teams Pursue *Varroa* Mite Controls. *FarmProgress*. January 22, 2021 from <https://www.farmprogress.com/usda-bee-lab-teams-pursue-varroa-mite-controls>
- Colony Collapse Disorder (2021, January). US Environmental Protection Agency. <https://www.epa.gov/pollinator-protection/colony-col>

lapse-disorder

Calovi, M., et al. Summer weather conditions influence winter survival of honey bees (*Apis mellifera*) in the northeastern United States. *Sci Rep* 11, 1553 (2021).
<https://doi.org/10.1038/s41598-021-81051-8>

Damiani N., et al. Acaricidal and insecticidal activity of essential oils on *Varroa destructor* (Acari: Varroidae) and *Apis mellifera* (Hymenoptera: Apidae). *Parasitol Res.* 2009 Dec;106(1):145-52.

Davis C., and Ward W. (2003, December). Control of Chalkbrood Disease with Natural Products. Australian Government Rural Industries Research and Development Corporation.
<https://www.agrifutures.com.au/wp-content/uploads/publications/03-107.pdf>

Garrido, C. (2018, February 27). Thymol *Varroa* Control. Bee Culture.
<https://www.beeeculture.com/thymol-varroa-control/>

Ghasemi V, Moharramipour S, and Tahmasbi G. (2011). Biological activity of some plant essential oils against *Varroa destructor* (Acari: Varroidae), an ectoparasitic mite of *Apis mellifera* (Hymenoptera: Apidae). *Exp Appl Acarol.* Oct;55(2):147-54.

Helping Agriculture's Helpful Honey bees. (2021, January). US Food and Drug Administration.
<https://www.fda.gov/animal-veterinary/animal-health-literacy/helping-agricultures-helpful-honey-bees>

Hunt G.J., and Given K. Beekeeping What Beekeepers Should Know About Bee Mites. Purdue University Extension.
<https://extension.entm.purdue.edu/publications/E-201/E-201.html>

Imdorf A., et al. (1999). Use of Essential Oils for the Control of *Varroa jacobsoni* Oud. in Honey Bee Colonies. *Apidologie, Springer Verlag*, 30 (2-3), pp.209-228.

Moore P.A., Wilson M.E., and Skinner J.A. (2019 August 20). Honey Bee Viruses, the Deadly *Varroa* Mite Associates. Bee Health, January 23, 2021 from <https://bee-health.extension.org/honey-bee-viruses-the-deadly-varroa-mite-associates/>

Noël A, Le Conte Y, and Mondet F. *Varroa*

destructor: how does it harm *Apis mellifera* honey bees and what can be done about it? *Emerg Top Life Sci.* 2020 Jul 2;4(1):45-57. doi: 10.1042/ETLS20190125. PMID: 32537655; PMCID: PMC7326341

Ramsey S.D., et al. (2019). *Varroa destructor* feeds primarily on honey bee fat body tissue and not hemolymph. *Proceedings of the National Academy of Sciences.* 116 (5) 1792-1801.

Randall, B. (2020, June 22). The Value of Birds and Bees from USDA.
<https://www.farmers.gov/blog/conservation/value-birds-and-bees>

Rosenkranz P., Aumeier P., and Ziegelmann B. (2010) Biology and control of *Varroa destructor*. *J Invertebr Pathol* 103(Suppl 1):S96–S119.

Sabahi Q., et al. (2017). Continuous release of oregano oil effectively and safely controls *Varroa destructor* infestations in honey bee colonies in a northern climate. *Exp Appl Acarol.*72:263-275.

Stoner, A., Moffett J.O., and Wardecker, A.L. (1979). The Effect of the Todd Dead Bee Trap, Used Alone and in Combination with the Wardecker Waterer, on Colonies of Honey Bees (Hymenoptera: Apidae). *J of the Kansas Ent Soc*, vol. 52, no. 3, Kansas (Central States) Entomological Society, pp. 556–60.
<http://www.jstor.org/stable/25083950>.

Traynor, K., et al. (2016). In-hive Pesticide Exposome: Assessing risks to migratory honey bees from in-hive pesticide contamination in the Eastern United States. *Sci Rep* 6, 33207.

Three Primary Essential Oils are Useful in Beekeeping. (2021). Tennessee's Honey Bees.
https://www.tennesseehoneybees.com/index.php?main_page=page&id=27&chapter=0

Tools for *Varroa* Management. (2016, February 16). Honey Bee Health Coalition. January 23, 2021 from https://honeybeehealthcoalition.org/wp-content/uploads/2016/03/HBHC-Guide_Varroa_Interactive_BW_23FEB2016.pdf

van Dooremalen C, Gerritsen L, and Cornelissen B, et al. (2012). Winter survival of individual honey bees and honey bee colonies depends on level of *Varroa destructor* infestation. *PLoS ONE* 7:e36285

Lifestyle Factors that Impact Progression of Alzheimer's Disease

Medha Sarkar¹*

¹Los Gatos High School, Los Gatos, CA USA

Received June 11, 2022; Revised December 27, 2022; Accepted, January 9, 2023

Abstract

While there have been decades of research looking into the causes of Alzheimer's disease (AD), the true underlying pathogenesis continues to be a focus of active mechanistic studies. Even so, there is little holistic understanding of what processes or therapies help alleviate the decline of memory and cognition. As a result, it is difficult for caregivers and healthcare providers to have tangible actions to focus on that are validated to maintain or improve an Alzheimer's patient's function and quality of life. Here, we present a literature review of the currently available research to examine which social and physiological factors, such as music, exercise, diet, and more, affect the disease progression and, of those, which slow it. We then summarize the currently available treatment options for Alzheimer's dementia based on the previous findings. For example, melatonin improves disrupted circadian rhythms, while aerobic exercise and Mediterranean diets give neurons protection against beta-amyloid dysfunction through increased brain derived neurotrophic factor (BDNF). Additionally, SSRI drugs can help maintain brain mass while music therapy can temporarily activate preserved brain regions that enable positive moods.

Keywords: Alzheimer's Disease, Dementia, Caregivers, Cognition

1. Introduction

Alzheimer's disease (AD) is one of the most prevalent neurodegenerative disorders, affecting 5.8 million people in the US (Matthews et al., 2019), according to the CDC, and more than 50 million worldwide (Hebert et al., 2013). In 2020, treatment costs added up to \$300,000,000,000 in the US and is thought to increase substantially as the disease progresses in patients with AD (Wong, 2020). Additionally, about 4% of the population in the US acts as an unpaid caretaker to someone with AD ("Caregiving for a Person with Alzheimer's Disease or a Related Dementia," 2019).

While the cause of the disease is still widely disputed, one of the main physiological characteristics is the buildup of beta-amyloid (plaques) and the tau protein (tangles) in the brain

(Gulisano et al., 2018; Mena et al., 1995; Seeman & Seeman, 2011; Takahashi et al., 2017). In tandem, there is disruption of neural networks, resulting in memory loss and cognitive decline. The buildup also affects glial cells, which act as the protector cells of the brain. The glial cells help to maintain homeostasis and provide support and protection to the neurons. The plaques and tangles have been associated with glial cell destruction, which then results in brain atrophy. Comparison of an unaffected brain versus one with Alzheimer's shows a significant decrease in volume and mass, particularly in the hippocampal regions. As a result, many brain systems are interrupted leading to the symptoms of memory loss and cognitive decline. Often, people with the disease rely on others to help facilitate daily life.

With the lack of knowledge on an optimal treatment or therapy for AD, caregivers often feel

* Corresponding Author
medhasarkar25@gmail.com

Advisor: Sravani Kondapavulur
sravani.kondapavulur@ucsf.edu

overwhelmed. 80% of people with Alzheimer's receive in-home care, and some studies support the idea that caregiving is important to treatment of AD (Vernon et al., 2019), however there are very few studies done on that topic. There are social and physiological factors that have been scientifically shown to help outcomes of disease progression for AD. We chose these factors based on the idea that each is a highly variable process that has been shown to affect an individual's experience with AD based on the way they experience it.

The objective of the review is to summarize from available literature on how social and physiological factors impact the progression of AD and explore currently available treatment options based on the findings. Our hypothesis is that awareness of these factors is important so caregivers can either minimize the adverse effects or maximize the potential therapeutic benefit to deal with a non-reversible condition (Iqbal et al., 2010). Additionally, viable treatment options can be sourced from these factors by sourcing the beneficial aspects and regulating them into a therapy or treatment. For example, in the realm of sleep, caregivers can administer melatonin to a patient to improve disrupted circadian rhythms while aerobic exercise and Mediterranean diets improves production of BDNF, which gives neurons protection against beta-amyloid dysfunction. In patients with a history of depression, SSRI drugs can help maintain brain mass which normally decreases in AD. Finally, music therapy could be used as adjunct therapy for patients who are in the late stages of the disease with confusion or depression. This review was limited by a variety of factors such as small sample sizes and lack of uniformity in metrics.

2. Social and physiological factors that affect disease progression, cognition, and quality of life

In this section, we explore the social and physiological factors in greater detail with emphasis on prior studies and their impact on the progression of the disease. Those covered are sleep, exercise, diet, mental health, and exposure to music, each of which has been studied in relation to the advancement of the disease and are extremely mutable processes.

2.1 Sleep

In Alzheimer's disease, sleep disturbance is common, due to both environmental and genetic factors. Multiple studies have explored the relationship between sleep quality and functional independence. In particular, studies have demonstrated that greater independent activities of daily living (IADL) function were associated with higher objective sleep efficiency (OSE) and a higher Pittsburgh Sleep Quality Index (PSQI), (Hodgson et al., 2014; McCurry et al., 2006; Zhou et al., 2019). Furthermore, Zhou et al. (2019) found that people who had a better sleep quality also were more likely to have better cognitive function, while Hodgson et al (2014) found that people who had better PSQI had improved quality of life with regards to the categories of "Feelings" (which encompass self- concept) and "Everyday life" (which describes the patient's ability to independently navigate their life). Interestingly, research in this field has also demonstrated a positive relationship between daytime sleep and severity of cognitive decline, including results from other studies (Bonanni et al., 2005; Ohadinia et al., 2004). However, sleep disturbance doesn't correlate with AD severity in all studies, indicating that genetic susceptibility may play a role in these heterogeneous results (Fernandez-Martinez et al., 2002). In conclusion, sleep in Alzheimer's dementia is an important factor that can regulate a patient's quality of life, particularly with regards to their independent function and other aspects of their social lives.

Many medications have been introduced to enhance sleep in Alzheimer's dementia. A negative relationship between melatonin and amyloid effects have been found in many studies (Cardinali et al., 2005; Lin et al., 2013; Pandi-Perumal et al., 2005). Interestingly, melatonin decreased sleep latency significantly relative to placebo treatment (Cruz-Aguilar et al., 2018). Further research would have to be done to see if it helped with cognitive decline. Additionally, antipsychotics are prescribed for the behavioral manifestations of AD which have been effective as a last resort; however, they can increase the risks of falls and cardiac side effects (Katz et al., 1999; Rocca et al., 2007; Rocha et al., 2006). Serotonergic agents did not show benefits to

cognitive deterioration in AD (Atri et al., 2018; D. A. Bennett, 2018; Severino et al., 2018). Both norepinephrine and formoterol are commonly available medications that are now in preclinical and clinical trials, with future results to be released regarding efficacy (Dang et al., 2014). Donepezil has many varying results on clinical efficacy with many inconsistencies, but one generally agreed upon finding is that it improves REM sleep percentage (Cooke et al., 2006; dos Santos Moraes et al., 2006; Naharci et al., 2015). According to those same studies, galantamine has had little to no benefit when it comes to improvement of sleep architecture; many studies so far demonstrating efficacy have been non-randomized trials, therefore future randomized controlled trials (RCTs) are needed to probe this therapy further. Interestingly, rivastigmine, despite being a popular treatment for AD, has not been studied in a large enough capacity to determine the exact effect, although it has been suggested that it decreases PSQI insignificantly. A non-medication option is bright light therapy (BLT), which various studies showed can be helpful to regulate sleep schedules and has no serious adverse effects (Ancoli-Israel et al., 2003; Fetveit & Bjorvatn, 2005). In conclusion, these medications have been beneficial in improving sleep quality, independent functioning, and quality of life; however, there is still little to no evidence that these therapies prevent disease progression.

2.2 Exercise

Many studies have validated that exercise does help in prevention of AD and help stagnate the progression of AD. Exercise does not affect cognitive decline in AD; the study that supported this opinion was not randomized so further research will have to be done to fully support it (Borges-Machado et al., 2019). Additionally, quality of life was unaffected as a result of the intervention. Exercise also produces necessary proteins to protect against AD, specifically brain-derived neurotrophic factor (BDNF). Many studies cite a negative correlation between BDNF levels and the progression of AD (Connor et al., 1997; Holback et al., 2005; Holsinger et al., 2000; Nigam et al., 2017; Peng et al., 2005; Phillips et al.,

1991). BDNF also acts as a catalyst for antioxidant enzymes, so a lack of this factor could predispose a neuron to beta-amyloid dysfunction. Wrann et al. (2013) also supports this claim while also detailing that exercise specifically induces the PGC-1 α /FNDC5 pathway, which can increase BDNF production. Additionally, in a simulation with AD mice, those with increased exercise were found to have less neurogenesis impairment, due to a higher number of BrdU/NeuN+ cells which is theorized to help stop the progression of AD. Of note, this work has not yet been replicated in humans (D. Kim et al., 2019). In summary, BDNF production is why exercise is important in the regulation of AD and BDNF itself could possibly be further researched as a possible treatment option for AD.

Aerobic exercise has been shown to be significantly effective as an adjunct treatment for AD. In a study done by Sampaio et al. with AD patients with a neurocognitive disorder and other health issues, increased aerobic capacity resulted in the greatest improvement of cognition (Sampaio et al., 2020). Adversely, lower body flexibility, agility, and dynamic balance were not associated with cognition at all. Additionally, many studies suggest aerobic exercise improves executive control, working memory, visuospatial memory, retention and reaction time in humans (Chaddock et al., 2010; Churchill et al., 2002; Herting & Nagel, 2012; Hillman et al., 2008; Skriver et al., 2014; Spirduso & Clifford, 1978; van Praag et al., 1999). Similarly, in a study by Vidoni et al. (2019), aerobic exercise was shown to be significantly better compared to stretching and toning on the Disability Assessment for Dementia (DAD), showing more stability for functional independence and increasing IADL and ADL independence. Additionally, oxygen uptake increased in an intervention group who performed aerobic exercise regularly in Sobol et al. by 13%; however, in this study there was no additional mechanistic experiment to determine whether increased oxygenation improved AD quality of life (Sobol et al., 2018). Finally, in a study by Liu et al. (2013), a mouse model of Alzheimer's, treadmill workouts for five months decreased the progression of plaque buildup in their brains and also soothed behavioral problems. Further research will have to be done to

see if these same results of exercise improving AD pathophysiology can be replicated in humans. To summarize, aerobic activity tends to be far superior to other forms of exercise and could be used as a preventative option for patients who still have the capability of motion or patients who are still in the early stages of the disease.

2.3 Diet

Diet is an important factor in preventing AD and also slowing the progression of AD. A number of factors could be at play, including vitamin deficiencies, ratios of macronutrients, and genetic contribution to diet response. One prospective cohort study found that lower concentrations of vitamin B12 and folate (B9) were associated with a two-fold increase in likelihood to develop AD compared with higher concentrations (Wang et al., 2001). Of note, while there is currently limited mechanistic information as to how B12 and folate deficiencies can contribute to AD, it has been well-established that B12 deficiency leads to homocysteine buildup, which is neurotoxic (Bhatia & Singh, 2015; Obeid & Herrmann, 2006). In addition, vitamin D3 deficiency has been inversely correlated with the risk of AD, as it increases phagocytosis of beta-amyloid. In rodents, a ketogenic diet increased the progression of beta-amyloid build-up, whereas intermittent fasting decreased this buildup, compared to a control group of transgenic rats with AD (Ito et al., 2011; Singh et al., 2014). A Mediterranean diet, which consists of olive oil, grains, and vegetables, in unaffected individuals was shown to reduce the chance of having AD by up to 40% (Abuznait et al., 2013; Daccache et al., 2011; Miquel et al., 2018; Scarmeas et al., 2006; Szczechowiak et al., 2019; A. Wu et al., 2004). Also, the Mediterranean diet can be paired with exercise to increase BDNF production, which negates the progression of AD. Finally, in a study done by Luchsinger et al. (2002), people who were in the highest percentile for consumption of fats and calories who also possessed the APOE 4 allele, an allele that has been well-associated with AD, were found to have increased risk of AD. Overall, diet should be carefully examined in individuals with a family history of AD and in the early stages of AD to

ensure adequate nutritional balance.

2.4 Depression

Major depressive disorder, commonly referred to as depression, has been thoroughly researched in recent years as a response to increased publicity of mental health. In a case-control study of 125 people, episodes of depression, particularly those in older adults, were positively correlated with a higher risk of Alzheimer's disease/dementia compared to non-depressed individuals (Cantón-Habas et al., 2020). Additional research suggests the reason behind this trend, with protein metabolism being correlated with symptoms of depression (K.-Y. Wu et al., 2018; Zvěřová et al., 2013). Cortisol levels have also been suggested to be associated with cognitive decline in people with AD and symptoms of depression. Depression and dementia are also theorized to have similar pathophysiologic pathways; an alternate hypothesis is that depression is a prodromal symptom of dementia (Amieva et al., 2008; S. Bennett & Thomas, 2014; Wilson et al., 2008). Overall, depression is more than likely a major risk factor for AD and dementia, making it important to be aware of by caregivers.

Different depression treatments have been shown to help slow the progression of AD. Selective serotonin reuptake inhibitors (SSRI) are drugs that stop the reabsorption of the neurotransmitter serotonin back into the presynaptic neuron. Fluoxetine, commercially known as Prozac, is an SSRI drug that has been associated with treatment of AD. In many studies, use of fluoxetine has been positively associated with the mass of the hippocampus, hypothalamus, and anterior cortex (Bath et al., 2012; Chen et al., 2007; Imoto et al., 2015; Jin et al., 2017; Mendez-David et al., 2014; Phillips et al., 1991; Reynolds et al., 1995; Santarelli et al., 2003). Interestingly, in Alzheimer's disease there is atrophy of the frontal and temporal cortices. Thus, hippocampal mass preservation, in the temporal cortex, could demonstrate a structural basis of therapeutic effect of antidepressants in the therapeutic efficacy for both depression and Alzheimer's disease. BDNF levels have also been positively correlated with fluoxetine, providing

evidence towards additional neuropharmacological benefits. Other SSRI drugs, including sertraline, citalopram, trazodone, and moclobemide, have also been similarly effective. Buspirone, and mirtazapine have not been studied enough to come to a conclusion (Buhr & White, 2007; Herrmann & Lanctôt, 2007; Lyketsos et al., 2003; Sink et al., 2005; Starkstein & Mizrahi, 2006).

Some studies have found that while psychotherapy is effective initially, therapeutic benefit decreases over time; thus they recommend psychotherapy as a short-term option for comorbid depression and Alzheimer's disease. However, other studies demonstrate long-term efficacy in psychotherapy treatment of depression in the elderly, thus further large-cohort studies are needed to demonstrate true effect (Farina et al., 2017; Leong, 2014). In the study by Wilkins et al. (2010), it was especially found that efficacy of cognitive behavioral therapy (CBT) was dependent on note-taking and frequent summaries in patients with dementia to account for memory loss, in order for therapy to be effective. This perhaps could have contributed to the decreasing efficacy seen in other studies. In conclusion, SSRI therapy can be considered a first-line therapy for patients with comorbid Alzheimer's and depression, but psychotherapy could also be a beneficial second-line therapy.

2.5 Music

Music therapy has been researched as a short-term therapy for AD patients. In general, long-term music-related memory remains more preserved, whereas short-term and autobiographical long-term memory are relatively more affected in patients (Dahms et al., 2021). The latter two are represented in the hippocampus, which is very quickly atrophied in AD. As a result, patients with AD show the most intense response to music that they prefer and are familiar with, as it activates the supplementary motor area of the brain that is not affected in the early progression of AD. Additionally, emotionally charged music was shown to recall autobiographical memory efficiently, especially events that were emotional or important to the patient (Guetin et al., 2013). Music therapy also slows the decrease in Mini-Mental State Exams

(MMSE) scores. In one study, MMSE scores of a patient group who had received music therapy was significantly better than the control group. In a study done in China, Mozart and Liangzhu music, ancient Chinese orchestral music, were shown to be the most effective at improving quality of life indices. This may be because the population studied was an elderly Chinese population and therefore could have been exposed to this music in their childhood and adolescence. Indeed, Arroyo-Anlló et al. (2013) demonstrated that while patients with AD experienced a decline in MMSE scores with unfamiliar music exposure, those who were exposed to familiar music had no change in MMSE scores. In a case study from Japan, while music therapy improved a client's social interaction, it did not affect the client's depression and progression of dementia. Finally, other studies have also linked music therapy with increased global cognition, decreased anxiety and irritability, and interestingly improved quality of life of caregivers (Bruer et al., 2007; Gómez Gallego & Gómez García, 2017; Herholz et al., 2013; H.-J. Kim et al., 2016). Overall, while there is no strong evidence that music therapy slows the progression of AD pathology, it can be used as a method for improving quality of life for both AD patients and caregivers as well as caregivers, although there is no strong evidence towards slowing the progression of the disease pathology.

3. Discussion

This review highlights some of the keyways that caregivers can approach helping patients with Alzheimer's disease (AD). We found many ways to improve quality of life in a patient and slow the progression of the disease. For example, in the realm of sleep, caregivers can administer melatonin to a patient to improve disrupted circadian rhythms or aid with falling asleep as it decreases sleep latency. Aerobic exercise should be encouraged in patients who still have full range of movement as it improves production of BDNF, which gives neurons protection against beta-amyloid dysfunction. Additionally, a Mediterranean diet also is correlated with increased BDNF production while also decreasing the risk of AD by 40%, which can be incorporated into a

patient's daily meals. In patients with a history of depression, SSRI drugs can not only help with mood symptoms, but also help maintain brain mass which normally decreases in AD. Finally, music therapy could be used as adjunct therapy for patients who are in the late stages of the disease with confusion or depression, due to temporary activation of preserved brain regions that enable positive mood and music-related associations.

This review was limited by a variety of factors. For example, the studies examining the role of sleep in AD involved many different metrics for sleep quality, making it difficult to directly compare the results. In many of the studies, rodents were used as the animal model. While rodent brains are useful in simulating many human processes, it is still hard to validate the results as rodent brains still have many significant differences from a human brain. The sample sizes were sufficient in most of the papers, however some, especially those examining the relationship between AD and depression, were not big enough to fully validate the results. While we initially were looking to characterize how best to support patients with late-stage AD, we found that most of the studies, including the ones here, examined effects on patients with early-stage AD. This could be because late-stage AD patients may not be able to participate in a study to the full extent as compared to a patient who has just been diagnosed. Further studies will need to be conducted to further examine whether these lifestyle protective measures also can be therapeutic in late-stage AD.

4. Conclusion

This paper focuses on therapies that are non-invasive and interventions that caregivers can easily administer. The results of this study can be used to facilitate home life of AD patients, as well as make it easier for them to deal with the debilitating effects of the disease in a way that improves quality of life. In particular, we examined the roles of sleep, diet, exercise, antidepressant medication, and music in improving AD symptomatology. We hope that the summarized results provide new tangible ideas for caregivers to help implement to improve the daily lives of patients with AD.

Acknowledgment

I would like to thank my parents, my grandfather, my grandmother, and my friends for their continued support of my interests and my endeavors.

References

- Abuznait, A. H., et al. (2013). Olive-Oil-Derived Oleocanthal Enhances β -Amyloid Clearance as a Potential Neuroprotective Mechanism against Alzheimer's Disease: In Vitro and in Vivo Studies. *ACS Chemical Neuroscience*, 4(6), 973–982. <https://doi.org/10.1021/cn400024q>
- Amieva, H., et al. (2008). Prodromal Alzheimer's disease: Successive emergence of the clinical symptoms. *Annals of Neurology*, 64(5), 492–498. <https://doi.org/10.1002/ana.21509>
- Ancoli-Israel, S., et al. (2003). Increased Light Exposure Consolidates Sleep and Strengthens Circadian Rhythms in Severe Alzheimer's Disease Patients. *Behavioral Sleep Medicine*, 1(1), 22–36. https://doi.org/10.1207/S15402010BSM0101_4
- Arroyo-Anlló, E. M., Díaz, J. P., & Gil, R. (2013). Familiar Music as an Enhancer of Self-Consciousness in Patients with Alzheimer's Disease. *BioMed Research International*, 2013, e752965. <https://doi.org/10.1155/2013/752965>
- Atri, A., Frölich, et al. (2018). Effect of Idalopirdine as Adjunct to Cholinesterase Inhibitors on Change in Cognition in Patients With Alzheimer Disease: Three Randomized Clinical Trials. *JAMA*, 319(2), 130–142. <https://doi.org/10.1001/jama.2017.20373>
- Bath, K. G., et al. (2012). BDNF Val66Met Impairs Fluoxetine-Induced Enhancement of Adult Hippocampus Plasticity. *Neuropsychopharmacology*, 37(5), Article 5. <https://doi.org/10.1038/npp.2011.318>
- Bennett, D. A. (2018). Lack of Benefit With Idalopirdine for Alzheimer Disease: Another Therapeutic Failure in a Complex Disease Process. *JAMA*, 319(2), 123–125. <https://doi.org/10.1001/jama.2017.19700>
- Bennett, S., & Thomas, A. J. (2014). Depression and dementia: Cause, consequence or coincidence? *Maturitas*, 79(2), 184–190.

<https://doi.org/10.1016/j.maturitas.2014.05.009>

Bhatia, P., & Singh, N. (2015). Homocysteine excess: Delineating the possible mechanism of neurotoxicity and depression. *Fundamental & Clinical Pharmacology*, 29(6), 522–528. <https://doi.org/10.1111/fcp.12145>

Bonanni, E., et al. (2005). Daytime sleepiness in mild and moderate Alzheimer’s disease and its relationship with cognitive impairment. *Journal of Sleep Research*, 14(3), 311–317. <https://doi.org/10.1111/j.1365-2869.2005.00462.x>

Borges-Machado, F., et al.(2019). Feasibility and Impact of a Multicomponent Exercise Intervention in Patients With Alzheimer’s Disease: A Pilot Study. *American Journal of Alzheimer’s Disease & Other Dementias®*, 34(2), 95–103. <https://doi.org/10.1177/1533317518813555>

Bruer, R. A., Spitznagel, E., & Cloninger, C. R. (2007). The Temporal Limits of Cognitive Change from Music Therapy in Elderly Persons with Dementia or Dementia-Like Cognitive nmpairment: A Randomized Controlled Trial. *Journal of Music Therapy*, 44(4), 308–328. <https://doi.org/10.1093/jmt/44.4.308>

Buhr, G. T., & White, H. K. (2007). Difficult Behaviors in Long-term Care Patients With Dementia. *Journal of the American Medical Directors Association*, 8(3, Supplement 2), e101–e113. <https://doi.org/10.1016/j.jamda.2006.12.012>

Cantón-Habas, V., et al. (2020). Depression as a Risk Factor for Dementia and Alzheimer’s Disease. *Biomedicines*, 8(11), Article 11. <https://doi.org/10.3390/biomedicines8110457>

Cardinali, D. P., Furio, A. M., & Reyes, M. P. (2005). Clinical Perspectives for the Use of Melatonin as a Chronobiotic and Cytoprotective Agent. *Annals of the New York Academy of Sciences*, 1057(1), 327–336. <https://doi.org/10.1196/annals.1356.025>

Caregiving for a Person with Alzheimer’s Disease or a Related Dementia. (2019). *Centers for Disease Control and Prevention*. <https://www.cdc.gov/aging/caregiving/alzheimer.htm#print>

Chaddock, L., et al. (2010). Basal Ganglia Volume Is Associated with Aerobic Fitness in Preadolescent Children. *Developmental Neuroscience*, 32(3), 249–256. <https://doi.org/10.1159/000316648>

Chen, S., et al.(2007). Serotonin stimulates mitochondrial transport in hippocampal neurons. *Molecular and Cellular Neuroscience*, 36(4), 472–483. <https://doi.org/10.1016/j.mcn.2007.08.004>

Churchill, J. D., et al. (2002). Exercise, experience and the aging brain. *Neurobiology of Aging*, 23(5), 941–955. [https://doi.org/10.1016/S0197-4580\(02\)00028-3](https://doi.org/10.1016/S0197-4580(02)00028-3)

Connor, B., et al.. (1997). Brain-derived neurotrophic factor is reduced in Alzheimer’s disease. *Molecular Brain Research*, 49(1), 71–81. [https://doi.org/10.1016/S0169-328X\(97\)00125-3](https://doi.org/10.1016/S0169-328X(97)00125-3)

Cooke, J. R., et al.(2006). Acetylcholinesterase Inhibitors and Sleep Architecture in Patients with Alzheimer’s Disease. *Drugs & Aging*, 23(6), 503–511. <https://doi.org/10.2165/00002512-200623060-00005>

Cruz-Aguilar, M. A., et al.(2018). Melatonin Effects on EEG Activity During Sleep Onset in Mild-to-Moderate Alzheimer’s Disease: A Pilot Study. *Journal of Alzheimer’s Disease Reports*, 2(1), 55–65. <https://doi.org/10.3233/ADR-170019>

Daccache, A., et al. (2011). Oleuropein and derivatives from olives as Tau aggregation inhibitors. *Neurochemistry International*, 58(6), 700–707. <https://doi.org/10.1016/j.neuint.2011.02.010>

Dahms, R., et al. (2021). Influence of Music Therapy and Music-Based Interventions on Dementia: A Pilot Study. *Journal of Music Therapy*, 58(3), e12–e36. <https://doi.org/10.1093/jmt/thab005>

Dang, V., et al. (2014). Formoterol, a Long-Acting β 2 Adrenergic Agonist, Improves Cognitive Function and Promotes Dendritic Complexity in a Mouse Model of Down Syndrome. *Biological Psychiatry*, 75(3), 179–188. <https://doi.org/10.1016/j.biopsych.2013.05.024>

dos Santos Moraes, W. A., et al. (2006). The Effect of Donepezil on Sleep and REM Sleep EEG in Patients with Alzheimer Disease: A Double- Blind Placebo-Controlled Study. *Sleep*, 29(2), 199–205.

<https://doi.org/10.1093/sleep/29.2.199>

Farina, N., Morrell, L., & Banerjee, S. (2017). What is the therapeutic value of antidepressants in dementia? A narrative review. *International Journal of Geriatric Psychiatry*, 32(1), 32–49. <https://doi.org/10.1002/gps.4566>

Fernandez-Martinez, M., et al. (2002). Prevalence of Neuropsychiatric Symptoms in Alzheimers Disease and Vascular Dementia. *Current Alzheimer Research*, 5(1), 61–69.

Fetveit, A., & Bjorvatn, B. (2005). Bright-Light Treatment Reduces Actigraphic-Measured Daytime Sleep in Nursing Home Patients With Dementia: A Pilot Study. *The American Journal of Geriatric Psychiatry*, 13(5), 420–423. <https://doi.org/10.1097/00019442-200505000-00012>

Gómez Gallego, M., & Gómez García, J. (2017). Music therapy and Alzheimer’s disease: Cognitive, psychological, and behavioural effects. *Neurología (English Edition)*, 32(5), 300–308. <https://doi.org/10.1016/j.nrleng.2015.12.001>

Guetin, S., et al. (2013). An overview of the use of music therapy in the context of Alzheimer’s disease: A report of a French expert group. *Dementia*, 12(5), 619–634. <https://doi.org/10.1177/1471301212438290>

Gulisano, W., et al. (2018). Role of Amyloid- β and Tau Proteins in Alzheimer’s Disease: Confuting the Amyloid Cascade. *Journal of Alzheimer’s Disease*, 64(s1), S611–S631. <https://doi.org/10.3233/JAD-179935>

Hebert, L. E., et al. (2013). Alzheimer disease in the United States (2010–2050) estimated using the 2010 census. *Neurology*, 80(19), 1778–1783. <https://doi.org/10.1212/WNL.0b013e31828726f5>

Herholz, S. C., Herholz, R. S., & Herholz, K. (2013). Non-pharmacological interventions and neuroplasticity in early stage Alzheimer’s disease. *Expert Review of Neurotherapeutics*, 13(11), 1235–1245. <https://doi.org/10.1586/14737175.2013.845086>

Herrmann, N., & Lanctôt, K. L. (2007). Pharmacologic Management of Neuropsychiatric Symptoms of Alzheimer Disease. *The Canadian Journal of Psychiatry*, 52(10), 630–646.

<https://doi.org/10.1177/070674370705201004>

Herting, M. M., & Nagel, B. J. (2012). Aerobic fitness relates to learning on a virtual Morris Water Task and hippocampal volume in adolescents. *Behavioural Brain Research*, 233(2), 517–525. <https://doi.org/10.1016/j.bbr.2012.05.012>

Hillman, C. H., Erickson, K. I., & Kramer, A. F. (2008). Be smart, exercise your heart: Exercise effects on brain and cognition. *Nature Reviews Neuroscience*, 9(1), Article 1. <https://doi.org/10.1038/nrn2298>

Hodgson, N., Gitlin, L. N., & Huang, J. (2014). The influence of sleep disruption and pain perception on indicators of quality of life in individuals living with dementia at home. *Geriatric Nursing*, 35(5), 394–398. <https://doi.org/10.1016/j.gerinurse.2014.08.005>

Holback, S., Adlerz, L., & Iverfeldt, K. (2005). Increased processing of APLP2 and APP with concomitant formation of APP intracellular domains in BDNF and retinoic acid-differentiated human neuroblastoma cells. *Journal of Neurochemistry*, 95(4), 1059–1068. <https://doi.org/10.1111/j.1471-4159.2005.03440.x>

Holsinger, R. M. D., et al. (2000). Quantitation of BDNF mRNA in human parietal cortex by competitive reverse transcription-polymerase chain reaction: Decreased levels in Alzheimer’s disease. *Molecular Brain Research*, 76(2), 347–354. [https://doi.org/10.1016/S0169-328X\(00\)00023-1](https://doi.org/10.1016/S0169-328X(00)00023-1)

Imoto, Y., et al. (2015). Role of the 5-HT4 receptor in chronic fluoxetine treatment-induced neurogenic activity and granule cell dematuration in the dentate gyrus. *Molecular Brain*, 8(1), 29. <https://doi.org/10.1186/s13041-015-0120-3>

Iqbal, K., et al. (2010). Tau in Alzheimer Disease and Related Tauopathies. *Current Alzheimer Research*, 7(8), 656–664. <https://doi.org/10.2174/156720510793611592>

Ito, S., Ohtsuki, et al. (2011). 1 α ,25-Dihydroxyvitamin D3 enhances cerebral clearance of human amyloid- β peptide(1-40) from mouse brain across the blood-brain barrier. *Fluids and Barriers of the CNS*, 8(1), 20. <https://doi.org/10.1186/2045-8118-8-20>

- Jin, H.-J., et al. (2017). Alleviative effects of fluoxetine on depressive-like behaviors by epigenetic regulation of BDNF gene transcription in mouse model of post-stroke depression. *Scientific Reports*, 7(1), Article 1. <https://doi.org/10.1038/s41598-017-13929-5>
- Katz, I. R., et al. (1999). Comparison of Risperidone and Placebo for Psychosis and Behavioral Disturbances Associated With Dementia: A Randomized, Double-Blind Trial. *The Journal of Clinical Psychiatry*, 60(2), 3507. <https://doi.org/10.4088/JCP.v60n0207>
- Kim, D., Cho, J., & Kang, H. (2019). Protective effect of exercise training against the progression of Alzheimer's disease in 3xTg-AD mice. *Behavioural Brain Research*, 374, 112105. <https://doi.org/10.1016/j.bbr.2019.112105>
- Kim, H.-J., et al. (2016). Effectiveness of a community-based multidomain cognitive intervention program in patients with Alzheimer's disease. *Geriatrics & Gerontology International*, 16(2), 191–199. <https://doi.org/10.1111/ggi.12453>
- Leong, C. (2014). Antidepressants for Depression in Patients with Dementia: A Review of the Literature. *The Consultant Pharmacist*, 29(4), 254–263. <https://doi.org/10.4140/TCP.n.2014.254>
- Lin, L., et al. (2013). Melatonin in Alzheimer's Disease. *International Journal of Molecular Sciences*, 14(7), Article 7. <https://doi.org/10.3390/ijms140714575>
- Liu, H., et al. (2013). Long-term treadmill exercise inhibits the progression of Alzheimer's disease-like neuropathology in the hippocampus of APP/PS1 transgenic mice. *Behavioural Brain Research*, 256, 261–272. <https://doi.org/10.1016/j.bbr.2013.08.008>
- Luchsinger, J. A., et al. (2002). Caloric Intake and the Risk of Alzheimer Disease. *Archives of Neurology*, 59(8), 1258–1263. <https://doi.org/10.1001/archneur.59.8.1258>
- Lyketsos, C. G., et al. (2003). Treating Depression in Alzheimer Disease: Efficacy and Safety of Sertraline Therapy, and the Benefits of Depression Reduction: The DIADS. *Archives of General Psychiatry*, 60(7), 737–746. <https://doi.org/10.1001/archpsyc.60.7.737>
- Matthews, K. A., et al. (2019). Racial and ethnic estimates of Alzheimer's disease and related dementias in the United States (2015–2060) in adults aged ≥ 65 years. *Alzheimer's & Dementia*, 15(1), 17–24. <https://doi.org/10.1016/j.jalz.2018.06.3063>
- McCurry, S. M., et al. (2006). Factors Associated With Caregiver Reports of Sleep Disturbances in Persons With Dementia. *The American Journal of Geriatric Psychiatry*, 14(2), 112–120. <https://doi.org/10.1097/01.JGP.0000192499.25940.da>
- Mena, R., et al. (1995). Monitoring pathological assembly of tau and β - amyloid proteins in Alzheimer's disease. *Acta Neuropathologica*, 89(1), 50–56. <https://doi.org/10.1007/BF00294259>
- Mendez-David, I., et al. (2014). Rapid Anxiolytic Effects of a 5-HT₄ Receptor Agonist Are Mediated by a Neurogenesis-Independent Mechanism. *Neuropsychopharmacology*, 39(6), Article 6. <https://doi.org/10.1038/npp.2013.332>
- Miquel, S., et al. (2018). Poor cognitive ageing: Vulnerabilities, mechanisms and the impact of nutritional interventions. *Ageing Research Reviews*, 42, 40–55. <https://doi.org/10.1016/j.arr.2017.12.004>
- Naharci, M. I., et al. (2015). Galantamine improves sleep quality in patients with dementia. *Acta Neurologica Belgica*, 115(4), 563–568. <https://doi.org/10.1007/s13760-015-0453-9>
- Nigam, S. M., et al. (2017). Exercise and BDNF reduce A β production by enhancing α -secretase processing of APP. *Journal of Neurochemistry*, 142(2), 286–296. <https://doi.org/10.1111/jnc.14034>
- Obeid, R., & Herrmann, W. (2006). Mechanisms of homocysteine neurotoxicity in neurodegenerative diseases with special reference to dementia. *FEBS Letters*, 580(13), 2994–3005. <https://doi.org/10.1016/j.febslet.2006.04.088>
- Ohadina, S., et al. (2004). Evaluation of Insomnia and Daytime Napping in Iranian Alzheimer Disease Patients: Relationship With Severity of Dementia and Comparison With Normal Adults. *The American Journal of Geriatric Psychiatry*, 12(5), 517–522. <https://doi.org/10.1097/00019442-200409000-00010>
- Pandi-Perumal, S. R., et al. (2005). Melatonin and

- sleep in aging population. *Experimental Gerontology*, 40(12), 911–925.
<https://doi.org/10.1016/j.exger.2005.08.009>
- Peng, S., et al. (2005). Precursor form of brain-derived neurotrophic factor and mature brain-derived neurotrophic factor are decreased in the pre-clinical stages of Alzheimer's disease. *Journal of Neurochemistry*, 93(6), 1412–1421.
<https://doi.org/10.1111/j.1471-4159.2005.03135.x>
- Phillips, H. S., et al. (1991). BDNF mRNA is decreased in the hippocampus of individuals with Alzheimer's disease. *Neuron*, 7(5), 695–702.
[https://doi.org/10.1016/0896-6273\(91\)90273-3](https://doi.org/10.1016/0896-6273(91)90273-3)
- Reynolds, G. p., et al. (1995). 5- Hydroxytryptamine (5-HT)₄ receptors in post mortem human brain tissue: Distribution, pharmacology and effects of neurodegenerative diseases. *British Journal of Pharmacology*, 114(5), 993–998.
<https://doi.org/10.1111/j.1476-5381.1995.tb13303.x>
- Rocca, P., et al. (2007). Risperidone, olanzapine and quetiapine in the treatment of behavioral and psychological symptoms in patients with Alzheimer's disease: Preliminary findings from a naturalistic, retrospective study. *Psychiatry and Clinical Neurosciences*, 61(6), 622–629.
<https://doi.org/10.1111/j.1440-1819.2007.01729.x>
- Rocha, F. L., et al. (2006). An Exploratory Open-Label Trial of Ziprasidone for the Treatment of Behavioral and Psychological Symptoms of Dementia. *Dementia and Geriatric Cognitive Disorders*, 22(5–6), 445–448.
<https://doi.org/10.1159/000095804>
- Sampaio, A., et al. (2020). Physical fitness in institutionalized older adults with dementia: Association with cognition, functional capacity and quality of life. *Aging Clinical and Experimental Research*, 32(11), 2329–2338.
<https://doi.org/10.1007/s40520-019-01445-7>
- Santarelli, L., et al. (2003). Requirement of Hippocampal Neurogenesis for the Behavioral Effects of Antidepressants. *Science*, 301(5634), 805–809. <https://doi.org/10.1126/science.1083328>
- Scarmeas, N., et al. (2006). Mediterranean diet and risk for Alzheimer's disease. *Annals of Neurology*, 59(6), 912–921. <https://doi.org/10.1002/ana.20854>
- Seeman, P., & Seeman, N. (2011). Alzheimer's disease: β -amyloid plaque formation in human brain. *Synapse*, 65(12), 1289–1297.
<https://doi.org/10.1002/syn.20957>
- Severino, M., et al. (2018). Established amyloid- β pathology is unaffected by chronic treatment with the selective serotonin reuptake inhibitor paroxetine. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, 4, 215–223.
<https://doi.org/10.1016/j.trci.2018.04.005>
- Singh, B., et al. (2014). Association of Mediterranean Diet with Mild Cognitive Impairment and Alzheimer's Disease: A Systematic Review and Meta-Analysis. *Journal of Alzheimer's Disease*, 39(2), 271–282. <https://doi.org/10.3233/JAD-130830>
- Sink, K. M., Holden, K. F., & Yaffe, K. (2005). Pharmacological Treatment of Neuropsychiatric Symptoms of Dementia: A Review of the Evidence. *JAMA*, 293(5), 596–608.
<https://doi.org/10.1001/jama.293.5.596>
- Skriver, K., et al. (2014). Acute exercise improves motor memory: Exploring potential biomarkers. *Neurobiology of Learning and Memory*, 116, 46–58.
<https://doi.org/10.1016/j.nlm.2014.08.004>
- Sobol, N. A., et al. (2018). Change in Fitness and the Relation to Change in Cognition and Neuropsychiatric Symptoms After Aerobic Exercise in Patients with Mild Alzheimer's Disease. *Journal of Alzheimer's Disease*, 65(1), 137–145.
<https://doi.org/10.3233/JAD-180253>
- Spirduso, W. W., & Clifford, P. (1978). Replication of Age and Physical Activity Effects on Reaction and Movement Time. *Journal of Gerontology*, 33(1), 26–30. <https://doi.org/10.1093/geronj/33.1.26>
- Starkstein, S. E., & Mizrahi, R. (2006). Depression in Alzheimer's disease. *Expert Review of Neurotherapeutics*, 6(6), 887–895.
<https://doi.org/10.1586/14737175.6.6.887>
- Szczechowiak, K., Diniz, B. S., & Leszek, J. (2019). Diet and Alzheimer's dementia – Nutritional approach to modulate inflammation. *Pharmacology Biochemistry and Behavior*, 184, 172743.
<https://doi.org/10.1016/j.pbb.2019.172743>

- Takahashi, R. H., Nagao, T., & Gouras, G. K. (2017). Plaque formation and the intraneuronal accumulation of β - amyloid in Alzheimer's disease. *Pathology International*, 67(4), 185–193. <https://doi.org/10.1111/pin.12520>
- van Praag, H., Kempermann, G., & Gage, F. H. (1999). Running increases cell proliferation and neurogenesis in the adult mouse dentate gyrus. *Nature Neuroscience*, 2(3), Article 3. <https://doi.org/10.1038/6368>
- Vernon, E. K., et al. (2019). Caregiver-Care Recipient Relationship Closeness is Associated With Neuropsychiatric Symptoms in Dementia. *The American Journal of Geriatric Psychiatry*, 27(4), 349–359. <https://doi.org/10.1016/j.jagp.2018.11.010>
- Vidoni, E. D., et al. (2019). Aerobic Exercise Sustains Performance of Instrumental Activities of Daily Living in Early-stage Alzheimer's Disease. *Journal of Geriatric Physical Therapy* (2001), 42(3), E129–E134. <https://doi.org/10.1519/JPT.0000000000000172>
- Wang, H.-X., et al. (2001). Vitamin B12 and folate in relation to the development of Alzheimer's disease. *Neurology*, 56(9), 1188–1194. <https://doi.org/10.1212/WNL.56.9.1188>
- Wilkins, V. M., Kiosses, D., & Ravdin, L. D. (2010). Late-life depression with comorbid cognitive impairment and disability: Nonpharmacological interventions. *Clinical Interventions in Aging*, 5, 323–331. <https://doi.org/10.2147/CIA.S9088>
- Wilson, R. S., et al. (2008). Change in Depressive Symptoms During the Prodromal Phase of Alzheimer Disease. *Archives of General Psychiatry*, 65(4), 439–445. <https://doi.org/10.1001/archpsyc.65.4.439>
- Wong, W. (2020). Economic burden of Alzheimer disease and managed care considerations. *The American Journal of Managed Care*, 26(8 Suppl), S177–S183. <https://doi.org/10.37765/ajmc.2020.88482>
- Wrann, C. D., et al. (2013). Exercise Induces Hippocampal BDNF through a PGC-1 α /FNDC5 Pathway. *Cell Metabolism*, 18(5), 649–659. <https://doi.org/10.1016/j.cmet.2013.09.008>
- Wu, A., Ying, Z., & Gomez-Pinilla, F. (2004). Dietary Omega-3 Fatty Acids Normalize BDNF Levels, Reduce Oxidative Damage, and Counteract Learning Disability after Traumatic Brain Injury in Rats. *Journal of Neurotrauma*, 21(10), 1457–1467. <https://doi.org/10.1089/neu.2004.21.1457>
- Wu, K.-Y., et al. (2018). Diversity of neurodegenerative pathophysiology in nondemented patients with major depressive disorder: Evidence of cerebral amyloidosis and hippocampal atrophy. *Brain and Behavior*, 8(7), e01016. <https://doi.org/10.1002/brb3.1016>
- Zhou, G., et al. (2019). High prevalence of sleep disorders and behavioral and psychological symptoms of dementia in late-onset Alzheimer disease. *Medicine*, 98(50), e18405. <https://doi.org/10.1097/MD.00000000000018405>
- Zvěřová, M., et al. (2013). Plasma cortisol in Alzheimer's disease with or without depressive symptoms. *Medical Science Monitor : International Medical Journal of Experimental and Clinical Research*, 19, 681–689. <https://doi.org/10.12659/MSM.889110>

The Effects of Colored Surgical Masks on Emotion Recognition and Perception in Adolescents

Phuong Bui¹ *

¹Christchurch School, Christchurch, VA USA

Received July 18, 2022; Revised December 26, 2022; Accepted, January 20, 2023

Abstract

Since the start of the COVID-19 pandemic, many studies have explored the effects of face masks on emotion recognition, yet no studies were conducted to explore the effect of different colored face masks on emotion recognition. This paper investigated the effect of four colors: red, black, green, and blue, compared to white, on perception of emotional intensity and recognition accuracy. 34 high schoolers were tested through a survey consisting of 120 images of models wearing surgical masks with photoshopped color stimuli. A significant effect was found in the colors red and green. Red enhances the perceived emotional intensity, and green confuses people into thinking the expression is disgust.

Keywords: Emotion Recognition, Colors, Face Masks, Perception

1. Introduction

The ability to read the facial expressions of others is essential to successful social interactions. Humans start to develop their ability to read emotions as young as four months of age (Barbera, et al., 1976). That ability continues to develop from late childhood through adulthood and can be changed depending on the age, gender, and pubertal status of a person (Thomas, et al., 2007; Lawrence, et al., 2015). However, this ability has been affected recently by the SARS-CoV-2 (COVID-19) pandemic. Since the beginning of the COVID-19 pandemic, face masks have been essential to reduce the spread of the virus (Li, et al., 2020). When the pandemic hit many countries around the world, people were required to wear masks outdoors and indoors. Because of this public health need, wearing masks has become a normal part of social interactions for many people. As it has been about 2 years since the World Health Organization (WHO) declared COVID-19 a global

pandemic, many types of masks have been sold on the market with various designs and colors.

1.1 Effects of Masking on Emotion Recognition.

Mask wearing habits have a negative impact on the ability to read facial expressions. As the COVID-19 pandemic affected people globally, humans have developed a reliance on masking to prevent the transmission of COVID. The pandemic has also changed humans' ways of processing emotions (Barrick, et al., 2021). In Barrick's study, scientists discovered that people who spent more time around masked faces relied more on the cues from the eyes to assess emotions, even if the person they observed did not wear face masks. However, according to Schurgin's (2014) finding, some regions of the face signal more information for a successful emotion recognition (Schurgin, et al., 2014). This leads to inaccuracies in reading emotions daily as selective facial expressions, such as joy and disgust,

* Corresponding Author
pee.ha.bee@gmail.com

Advisor: Derick Delloro
derick.delloro@hawken.edu

require more visual information from the mouth region.

Then how do people process emotions that depend on visual cues from the mouth region? When most facial visual information is obscured by a mask, one's intake of visual information becomes limited to the eye region, a region that does not provide sufficient information for an accurate emotional recognition. Carbon (2020) found out that accuracy of facial expression judgment is reduced when looking at masked individuals, except for fearful and neutral faces. In addition, subjects in a similar study of emotional intensity perceived all emotions expressed by masked individuals as less intense (Pazhoohi, et al., 2021).

1.2 Effects of color on emotion recognition

Limiting visual information is not the only factor that affects emotional recognition. Differences in the color of objects in a subject's visual field may play a role in emotion recognition as well. Previous studies have found that facial color, clothing color, and background color can modulate the perception of emotional expression (Minami, et al., 2018; Nakajima, et al., 2017; Wiedemann, et al., 2015). However, these studies only tested three colors in their experiments: red, gray, and blue. Red was found to increase perceived anger and aggression while blue and gray had no effect on emotional recognition (Minami, et al., 2018; Wiedemann, et al., 2015). The potentially synergistic effects of the face masks' colors on emotion recognition have not yet been researched.

The association between color red and emotion is proven to have an impact on facial expression recognition. Therefore, with a mask occluding more than half of one's face, red is more likely to influence humans' perception of an expression as more angry and aggressive. Similar results would potentially be shown in different colors, with direct correlations with their positive or negative emotional associations. Furthermore, influenced by the mask color, misinterpretations of an emotion would be more likely to occur, causing inaccuracies during the emotion recognition process. With that in mind, this study aims to investigate the effects of colored

surgical masks on the ability to accurately recognize emotions and emotional intensity.

2. Method

In this project, data was gathered in-person, through a Qualtrics form based on similar methodology done by Pazhoohi (2021).

2.2 Participants

A total of 34 subjects participated in this study. All participants were high-school students (17 men, 16 women, and 1 preferred not to say) between 14 to 18 years of age. All participants were students at a boarding school in Eastern Virginia. Exclusion criteria included diagnoses of color blindness, autism, Asperger's, and Alexithymia, as self-reported at the beginning of the experiment. 17 of the participants self-identified as White, 6 Asian, 6 Black, 4 Latinos, and 1 Turkish. All participants provided written informed consent and participants in the experiment received academic credit and were entered in a raffle for a \$20 Amazon gift card. The study was approved by the Christchurch School (CCS) internal review board.

2.3 Stimuli

A total of 120 images were viewed by each participant. The images consisted of expressive faces posed by 2 female Caucasian adults and 2 male Caucasians adults taken from the FACES database (Ebner, et al., 2010). Each model provided six facial expressions: happiness, fear, anger, disgust, sadness, and neutrality. Surgical masks were photoshopped into the pictures using Adobe Photoshop, with 5 color stimuli in total: White, Black, Pure Blue #0025ff, Vivid Red #e20410, Strong Yellow (soft green) #a2bc00 (Figure 1). Faces with white masks are considered as the control group as white is the absence of color. Black, blue, and red were chosen according to a study that tested the effect of background color on facial expression perception (Minami, et al., 2018). The color green, however, was chosen based on its association with positive and powerful emotions such as relaxation and fear

(Jonauskaite, et al., 2020; Jonauskaite, et al., 2019).



Figure 1. Examples of masked facial expressions in order of anger, neutral, joy, disgust, and surprise. The original images were obtained from the MPI FACES database.

2.4 Procedure

Before beginning the experiment, participants were taken to a quiet room in the school’s science building. They were given a consent form to fill out and had their phones taken away. Collecting phones before starting the experiment was deemed necessary to maximize participant engagement during the study. Participants had the option to bring their own computer or use the one given to do the survey. The study was conducted in person to allow the researcher to monitor engagement and maintain a low distraction environment.

In the survey, each individual was shown randomly ordered images of various masked people making different facial expressions. Under each image, participants were asked to answer the following questions:

1. “What is the facial expression of this person?”
Their choices were: anger, disgust, fear, happy, neutral, fear and sadness.
2. “From 0 to 100%, how much of this emotion is the person expressing?” with a provided scale to interact. The question of expressive intensity was not asked for faces that were judged to be neutral.

Other than experimental questions, participants were also asked about demographic information such as their race, age, sex, and whether they meet the exclusionary criteria. The language used in this study was written at an appropriate level for students at an English-speaking school to understand. The experiment was conducted across a time span of two weeks.

2.4 Data processing and analysis methods

The data collected was analyzed using three statistical tests, the two-way ANOVA, Chi-squared, and *t*-test. Each test has a different objective: A two-way ANOVA test was performed on all colors to determine an overall effect of it on emotional intensity. As previous studies have established a clear effect of red on emotion recognition, an independent *t*-test was used to examine red and white separately. Finally, to determine whether color has an effect on emotion recognition accuracy, a Chi-squared test was conducted. A result is indicated as significant when its *p*-value is less than 0.05.

3. Results

3.1 Perceived intensity of facial expressions

Average emotional intensity ratings were compared across gender of models, gender of participants, and individually by each emotion. Emotions that were judged to be neutral do not have emotional intensity ratings. These ratings were pooled and compared by color of mask and no statistically significant data was found through two-way ANOVA tests. However, looking at the average emotion intensity ratings of all participants, models, and emotions pooled by color, a trend toward significance was found with a *p*-value of 0.0601 (Figure 2).

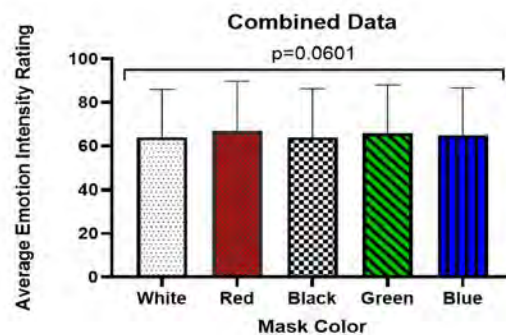


Figure 2. Average intensity of emotions by mask color. Ratings for emotional intensity were collected and averaged among all participants and compared by mask color of models. Data displayed in mean and SD. *p*=0.0601 via 2-way ANOVA test.

Intensity ratings from all participants looking at all models were averaged and compared for only white and red masks. The comparison was also performed across gender of models and gender of participant, although no significance was found. However, when looking at the combined ratings, the significant effect of red has on perceived intensity of all expressions was found through an independent t-test with a p-value of 0.0182 (Figure 3).

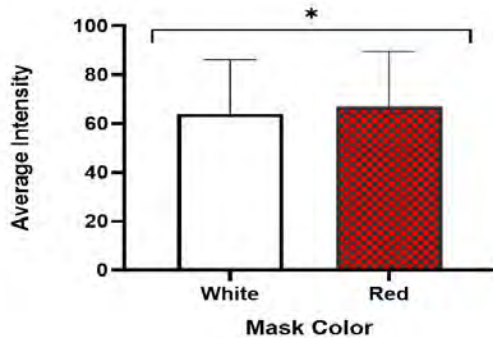


Figure 3. Effect of red on the perception of emotional intensity. Average emotional intensity ratings for all models and participants were compared for red and white masks only. Models wearing red masks received significantly higher emotional intensity ratings compared to the control group (white masks). Data displayed in mean and SD. $p < 0.05$ via independent t-test.

3.2 Emotion Recognition Accuracy.

All responses were sorted by color and broken into 5 categories: combined responses, responses from female participants, responses from male participants, responses toward facial expressions posed by female models, and responses toward facial expressions posed by male models. No significant differences were found in responses sorted by either male or female participants. However, results showed a significant effect of all colors on disgusted facial expressions with a noticeable difference between red and green masks ($\chi^2=12.9934$, $p < 0.05$; Figure 4A). The effect also appeared on disgusted facial expressions posed by both female models ($\chi^2=15.2795$, $p < 0.005$; Figure 4B) and male models ($\chi^2=9.6495$, $p < 0.05$; Figure 4C). Another significant effect was found in all participants' responses toward fearful facial expression posed by female models

($\chi^2=10.7489$, $p < 0.05$; Figure 4D).

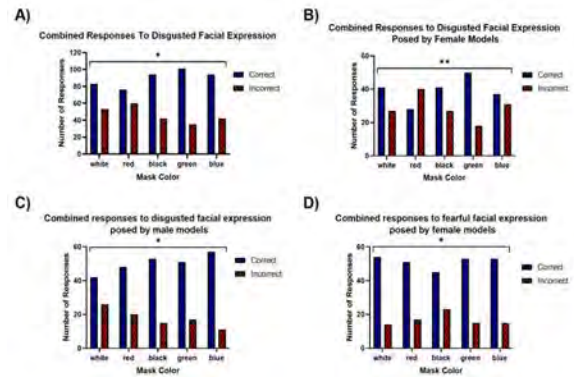


Figure 4. Emotional recognition accuracy by mask color. Correct and incorrect answers were counted and compared across different mask colors. Significant differences were observed in all responses to the disgusted facial expression (A), as well as in isolated responses to both female (B) and male (C) models making a disgusted facial expression. A significant difference was also observed in response to female models making a fearful facial expression (D). * $p < 0.05$, ** $p < 0.005$ via independent Chi-Square test.

4. Discussion and Limitations

In this study, the effect of colored surgical masks on emotion recognition was investigated by comparing all colors (red, black, green, blue) with white. Results showed that red does increase the perceived intensity of all emotions (Figure 3), which supports the previous findings of the color red's effect on clothing (Wiedemann, et al., 2015) and facial expressions (Nakajima, et al., 2017). Though the effect of the red mask on the perception of emotional intensity was found to be significant, results showed that the color only increased the perceived intensity to a small degree (Figure 3). Other than red, no significant effects of color on perceived emotional intensity were found.

Green was discovered to have a great effect on the ability to recognize fear and disgust. Participants' accuracy in recognizing disgusted facial expressions increased when models wore green face masks, whereas red face masks decreased all participants' accuracy in recognizing disgusted facial expressions (Figure 4A). However, this only occurred when the

red face mask was worn by female models. The specific mechanism of this gender-based difference should be explored through further research.

Not only does this investigation contribute to the known effects of red objects on perception of emotions and emotional intensity, but the results also suggest a potential effect of green. The contemporary relevance of these findings is sufficiently justified by the continuing SARS-CoV-2 pandemic. Humans are social beings, we socialize at work, at home, or at play and during this pandemic many people will continue to wear surgical masks daily. However, wearing a surgical mask obstructs parts of a face that are necessary for emotion recognition, and it's clear that mask color may influence the process as well by modulating the perception of emotional intensity and the accuracy of expression recognition. Therefore, the choices of mask color should be taken into consideration as the color chosen could enhance confusion and lead to unsuccessful social interactions.

In summary, red is the only color that increases the perceived intensity of all facial expressions (Fig. 3). The p-value found when testing the effect of all colors on modulation of intensity's perception is close to being statistically significant ($p=0.0601$). The similar study done by Pazhoohi (2021) had a much greater sample size-around 420 participants-compared to this study. If the participant pool were larger, the data may have shown more statistically significant effects of mask color on perception of emotions. Although this study had a limited sample size, several specific effects of mask color were observed. Green was discovered to increase participant's accuracy in the judgment of disgusted facial expressions: Even with its association with positive emotions (Jonaskaite, et al., 2020), participants were more likely to describe an expression as disgusted when the model was wearing a green mask. Green did not appear to affect the accuracy of recognizing emotions such as joy, anger, sadness, and neutral.

One potential limitation to this study is the lack of global diversity of participants. A total of 34 participants were recruited for this study, 6 of which were of Eastern Asian descent. Previous research has found that the interpretation of facial expressions

differs across cultures (Rachael, et al., 2009). According to the study, different cultures have different methods of identifying facial expressions: Eastern Asian people primarily fixate on the eye regions, whereas Westerners observe the whole face evenly. A person's method of recognizing facial expressions is uncontrollable, so this might have affected the results.

Additionally, the population pool only consisted of students from one international boarding school. Because of the ongoing pandemic at the time, recruiting participants from other schools was not a feasible task. This limits the accuracy of the result to reflect a larger population.

Finally, if the study only tested 3 colors, as opposed to 6, the result could turn out to be more accurate. Because the total number of face repetition in this study is 6 per expression, the similarity was easily noticed by participants. Moreover, it would reduce the surveying time from 40 minutes to 20 minutes. This helps increase participants' engagement overall.

5. Conclusion

To decrease the transmission rate of respiratory diseases, the practice of wearing masks indoors and outdoors is necessary. The usage of face masks has been shown to influence emotion recognition. According to Carbon (2020), wearing masks makes it harder for people to perceive genuine emotional states. For example, emotions such as disgust were frequently confused with anger. This study investigated the effects of colors on emotion recognition with existing confusion caused by face masks. The results strongly indicate that colors such as red and green affect emotion recognition accuracy. The perception of emotional intensity was also affected by color, specifically when a red surgical mask was worn. Models wearing green face masks were more likely to be accurately categorized as having a disgusted facial expression. Therefore, mask color was proven to be a factor that contributes to social interactions' success. However, the specific mechanisms of these color and gender-based effects need to be investigated through further research in order to be fully understood.

References

- Barrick, E. M., Thornton, M. A., & Tamir, D. I. (2021). Mask exposure during COVID-19 changes emotional face processing. *PLoS ONE*, [s. l.], 16(10), 1–22
- Carbon C-C (2020). Wearing Face Masks Strongly Confuses Counterparts in Reading Emotions. *Front. Psychol.* 11:566886. doi: 10.3389/fpsyg.2020.566886
- Ebner, N.C., Riediger, M., & Lindenberger, U. (2010). FACES—A database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavior Research Methods* 42, 351–362 (2010).
- Jonauskaite, D., et. al. (2020). Feeling Blue or Seeing Red? Similar Patterns of Emotion Associations With Colour Patches and Colour Terms. *i-Perception*, 11(1), 2041669520902484.
- Jonauskaite, D., et. al. (2019). What color do you feel? Color choices are driven by mood. *Color Research & Application*, 44(2), 272-284.
- La Barbera, J. D., et. al. (1976). Four- and six-month-old infants' visual responses to joy, anger, and neutral expressions. *Child Development*, 47, 535–538.
- Li, Y., et. al. (2020). Face masks to prevent transmission of COVID-19: A systematic review and meta-analysis. *American Journal of Infection Control*, 49(7), 900–906.
- Lawrence, K., Campbell, R., & Skuse, D. (2015). Age, gender, and puberty influence the development of facial emotion recognition. *Frontiers in psychology*, 6, 761.
- Minami, T., Nakajima, K. & Nakauchi, S. (2018) Effects of Face and Background Color on Facial Expression Perception. *Front. Psychol.* 9:1012.
- Nakajima, K., Minami, T., & Nakauchi, S. (2017). Interaction between facial expression and color. *Sci. Rep.* 7:41019. Doi: 10.1038/srep41019
- Pazhoohi, F., Forby, L., & Kingstone, A. (2021). Facial masks affect emotion recognition in the general population and individuals with autistic traits. *PLoS ONE*, 16(9), e0257740.
- Rachael. E. J., et. al. (2009). Cultural Confusions Show that Facial Expressions Are Not Universal. *Current Biology*, 19(18), 1543-1548, <https://doi.org/10.1016/j.cub.2009.07.051>
- Schurigin, M. W., et. al. (2014). Eye movements during emotion recognition in faces. *Journal of Vision*, 14(13): 14, 1-16, <https://doi.org/10.1167/14.13.14>
- Thomas, L. A., et. al. (2007). Development of emotional facial recognition in late childhood and adolescence. *Developmental science*, 10(5), 547–558.
- Wiedemann, D., et. al. (2015) Red clothing increases perceived dominance, aggression and anger. *Biol. Lett.* 11: 20150166. <http://dx.doi.org/10.1098/rsbl.2015.0166>

***In Silico* Drug Target Identification via Methylation Factors With Implementation in TP53 Liver Cancer Gene**

Sahithi Pogula^{1*}

¹Hopkinton High School, Hopkinton, MA USA

Received July 6, 2022; Revised January 13, 2023; Accepted, January 24, 2023

Abstract

The limitations of drug discovery are infamous, with a single drug development setting back an institution millions of dollars and decades of time with only a 0.1% rate of success. Novel identification of target leads for drugs through currently unused epigenetic measures reduces these hurdles, incredibly expediting the drug pipeline. Here we identified two novel approaches for reducing liver cancer. In the first approach, this project focuses on reducing the hypermethylation of 5-methylcytosine (5mC) by modulating the 5-hydroxymethylcytosine (5hmC) biological pathway. Using pattern search and similarity indexes, Algorithm-1 can identify potential positions of pathway induction by comparing aberrant sequences with normal sequences and using a CpG island pattern trigger. Algorithm-1 identified the TP53 liver cancer gene at positions 33-35. In the second approach, this project focuses on inhibiting spliceosome factors that are responsible for 5mC production. Using machine learning methods and pattern search, such spliceosome factors are identified. Algorithm-2 identified the Prp9 protein. Both the proposed solutions provide new insight into a cure for Liver Cancer never previously explored through traditional methods of drug discovery.

Keywords: Epigenomics, Liver Cancer, DNA Methylation, Spliceosomes, Drug

1. Introduction

Drug discovery in the present day is defined by an array of hurdles not limited to an extensive and inefficiently utilized amount of funds, years and even decades of research, as well as incredibly low Phase 1 and clinical trial approval rates. Characterized by traditional methods of research and thousands of diseases gone untreated, present-day drug discovery is something truly negatively impacting society. Among the diseases included in this category is Hepatocellular Carcinoma or Liver Cancer (Fan, et al., 2018). Combining the facts that HCC is the third-deadliest cancer in the world and that there is no present cure on the market (only peripheral solutions deterring side effects have been previously identified)

makes a deadly prediction of the future where the prevalence of HCC is only predicted to grow, coming as an economic, social, and moral burden to society. Epigenetics may possibly be the long-awaited solution to this highly sought-after problem.

Epigenetics refers to the study of heritable phenotype changes that do not involve alterations in the DNA sequence. Unlike genetic modifications, implications on epigenetics provide a novel tactic of gene expression modulation free of non-targeting side effects. DNA methylation, a naturally occurring epigenetic process, is closely correlated to embryonic development (Liu, et al., 2016), regulation of gene expression, X-chromosome inactivation, genomic imprinting, and genomic stability (Chen, et al., 2011). When located in the gene promoter, DNA

* Corresponding Author
sahithi.pogula@gmail.com

Advisor: Kristen Murphy
kmmurphy@hopkinton.k12.ma.us

methylation also acts to repress gene transcription. DNA methylation is catalyzed by a family of DNA methyltransferases (Dnmts) that transfer a methyl group from S-adenyl methionine (SAM) to the fifth carbon of a cytosine residue to form 5mC. Oftentimes, concentrated methylation sites are characterized as CpG sites (gene sites containing many C's and G's) (Chen, et al., 2011).

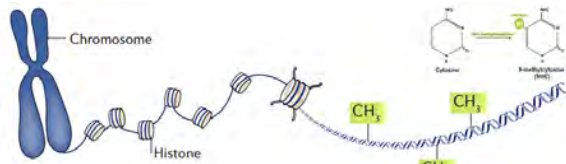


Figure 1. DNA methylation system naturally occurring via the utilization of the DNA methyltransferase (Wu, et al., 2017)

This natural process is made quite relevant to HCC studies due to its involvement in tumor progression. The development of HCC, in the current day, is marked by prevalent etiological conditions involving lifestyle as well as the environment. Recent studies, however, are coming to the surface stating otherwise (Hlady, et al. 2019). Aberrations in DNA methylation are now being directly correlated to molecular lesions of cancer cells as well as start points of tumor progression (Dreval, et al., 2019). The very regulated and maintained internal processes of DNA methylation was proven to be more radical and abnormal in HCC cases (Fan, et al., 2018). Hypermethylation was identified as a result.

Knowing the negative implications of such aberrant methylation, it is critical to identify these points of abnormality and revert, via demethylation, the hypermethylated cases of such points. The objective of this research was to accurately and efficiently identify loci of aberrant methylations for use in future applications of epigenetic modulation. Along with this, major motivations for this research included exploring new frontiers of epigenetic modulation, one being spliceosome factors and their inner entanglements to the onset of various carcinomas. Anticipated results for this research include studying the correlation between these various factors and methylation as well as building

algorithms and models relevant to the topic. A primary end goal of this research was to identify loci of aberrant methylation and specific spliceosome factors by using a computational model on Liver Cancer, particularly the TP52 oncogene

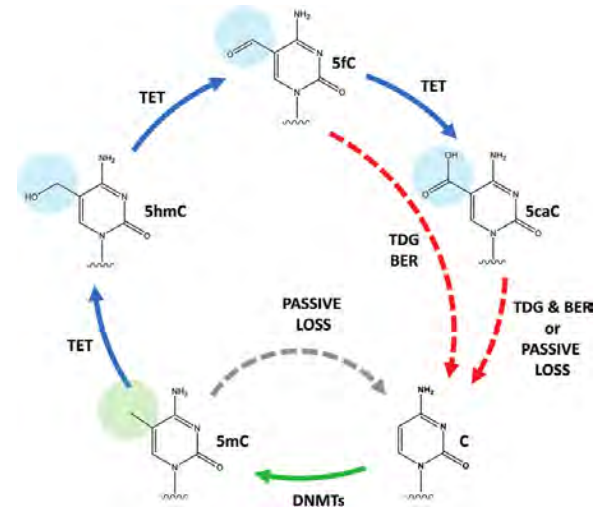


Figure 2. Multiple 5hmC DNA demethylation pathways. A methylated cytosine first starts its journey at a 5hmC (shown in the bottom left-hand corner). It is then oxidized by various TET family proteins to reach the state of 5caC. Through the BER process, it is then finally converted to an unmodified cytosine (shown in the bottom right-hand corner). (Chen, et al., 2011)

1.1 5hmC Pathway Demethylation Process

The 5-hydroxymethylcytosine (5hmC) pathway provides a viable solution for reducing such hypermethylation. As a whole, the 5hmC serves as a major role player in the roots of cancer development, not limited to simply Liver Cancer. In various types of carcinoma, aberrant methylation can be traced to be an underlying cause. Utilizing the 5hmC pathway, this aberrancy can be resolved. The exact points serving as tumor progressors oftentimes are characterized by extra methylation missing in a healthy case (Shi, et al., 2017). If these points are identified, the 5hmC pathway can be triggered artificially to revert methylation. The methyl moiety of 5mC is lost either passively during DNA replication or actively through enzymatic DNA demethylation. This project focuses on the latter. In a

process of oxidation reactions, conducted Ten-Eleven Translocation family proteins (TETs), the 5hmC is converted into a 5-carboxyl cytosine (5caC) (Dong, et al., 2012). Subsequently, the DNA base-excision repair (BER) pathway can also remove the methylated cytosine by filling in an unmodified cytosine. Thus, aberrancies can be reverted utilizing this pathway.

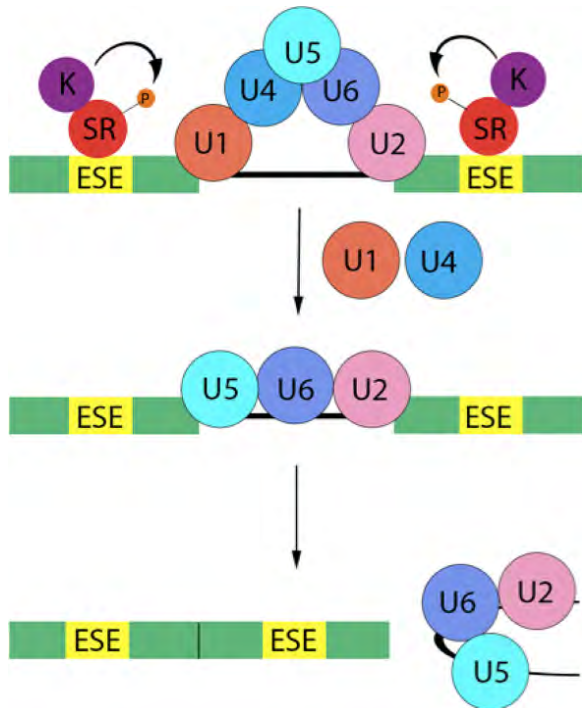


Figure 3. Alternative splicing mechanism as shown by spliceosome complex. Exonic splicing enhancers (ESEs) are present at the pre-mRNA splicing stage. The SR protein binds to ESE at the pre-mRNA splicing, but also are involved in mRNA export, genome stabilization, nonsense-mediated decay, and translation. (Wang, et al., 2015)

1.2 Spliceosome Factors

Modulating spliceosome behavior could also provide an effective solution to reducing the HCC-causing hypermethylation. This concept relies heavily on DNMT1-interacting RNAs (DiRs). DiRs are, essentially, the “blockers” between the DNMT1 and the gene. When a DiR is present, the DNMT1 is blocked from methylating at that site. In this way, DiRs regulate hypermethylation from occurring on genes. DiR use is not limited solely to Liver Cancer,

but also expands broadly to encompass the general methylation status of every gene in the body. Their behavior is most distinctively linked (as well as scientifically proven) to the onset of HCC and Liver Cancer. The production of DiRs is ensured through the spliceosome and the various spliceosome factors it contains to conduct alternative splicing (Wang, et al., 2015). When different spliceosome factors are used, splicing as well occurs differently—alternative splicing. In abnormal cases, DiRs are not spliced due to the spliceosome factors present. In this case, due to the repression of DiRs, the DNMT1 has unregulated access to the gene, resulting in hypermethylation. Once again, such hypermethylation serves as a lesion point to tumor progression. If the exact spliceosome factor responsible for DiR repression is identified, inhibition of it could revert the aberrant state of the gene. Splicing pattern analysis and matching individual factors to disease cases could reveal the etiological factor.

2. Materials and Methods

2.1 Phase 1

This phase comprises point-reversion of hypermethylation utilizing the 5-hmC pathway. Point reversion refers to reversing the methylation status of one singular base pair (in this case C) from a methylated to unmodified state. Through this phase, an algorithm is developed which is able to identify hypermethylation directly from the gene in mere minutes. Results then can be used as navigational points for 5-hmC pathway triggering.

To start off, various methylated and unmethylated gene sequences were collected from the public databases GenBank and KEGG. Previously identified sites of methylation were also retrieved from the application iMethyl. This application was built by many scientists that worked towards identifying methylations on healthy genes for more than a decade. Program enabling libraries were then set up—NumPy, Pandas, TensorFlow. A Naïve Brute Force algorithm was then used to sort through sequences. This method relies on a rather primitive tactic of more trial and error. Although it is quite time-taking, it guarantees accurate results without the

chance of any outside bias. In a taken methylated sequence, if a “substring” of it was identified to contain the pattern “CGCG”, the sequence is labeled. This “CGCG” pattern is characteristic of a CpG site where methylation is likely to present. In the sequence, if more than 3 substrings are identified to contain this pattern, methylation is evidently identified. Using Hamming Distance, a string similarity index, this methylated sequence is checked against an unmethylated sequence to detect erroneous methylation identification. The methylation sites identified are then checked against iMethyl’s methylation sites, using Hamming Distance, to identify hypermethylation.

2.2 Phase 2

This phase comprises spliceosome factor inhibition to prevent DiR repression. Through this phase, an algorithm is developed which is able to identify the spliceosome factor responsible for the specific splicing pattern that occurs in the case of DiR repression. Results can then be used to artificially inhibit the specific etiological factor identified.

Pre-mRNA and mRNA were first collected from the public database GenBank. Spliceosome edit patterns were also compiled from various NCBI publications. After ANN enabling programs, such as TensorFlow, were set up, basic tree nodes for a decision tree were built. The resulting decision tree was then trained multiple times using the data till cases of perfect fitting occurred. The Greedy Algorithm and Gini Index were used to quantify these accuracy rates. A Greedy Algorithm is a problem-solving heuristic that is able to find solutions to issues by inputting scenarios and evaluating them. This is important in accuracy rate calculations as it can serve as a comparison point for reconfirming solutions. The decision tree was then configured to output 3 top pattern matches for the splicing pattern seen in the pre-mRNAs, mRNAs, and the splicing factors. A Naïve Brute Force algorithm was then built to narrow down on the pattern match of the 3 top identified factors and the edit pattern. The end output of this algorithm is a splicing factor that is most likely to have acted upon the pre-mRNA

to convert it to the mRNA used.

2.3 Phase 3

Through this phase, the algorithms developed in the previous phases were used on the liver-cancer-associated TP53 gene. The Phase 1 algorithm identified hypermethylation at positions 33-35 on the TP53 gene. The Phase 2 algorithm found Prp5, Prp9, and Prp11 in the U2 unit of the spliceosome with the most pattern match to the TP53 gene. The Prp9 factor was then narrowed down as most probably responsible for DiR inhibition.

3. Results

In order to quantify error for Phase 1, the error distance from the methylation site identified and the methylation site present was calculated using the Jaccard Index. Trends are given below

3.1 Phase 1 algorithm summary of tests

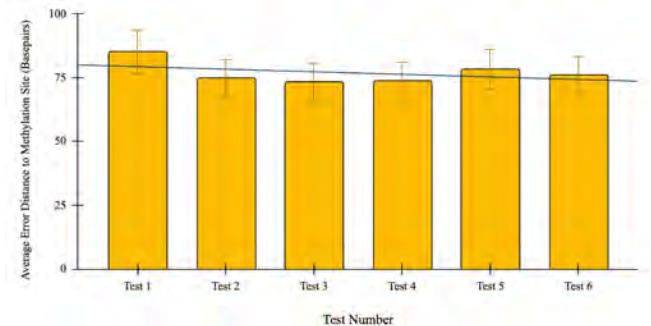


Figure 4. Average Error Distance to Methylation Site (set of 1,000 Basepairs)(units) vs. Test Number

As seen in the Phase 1 algorithm summary table, results stay relatively constant throughout out various tests and error distance remains low over the taken interval. These error distance values were used to calculate an 87.5% accuracy rate for the Phase 1 algorithm.

To test accuracy for the Phase 2 algorithm, the Greedy algorithm was used to assign values to how accurately a spliceosome factor was assigned to a given splicing pattern. Trends, given below, vary over tests due to the underfitting-overfitting aspect of machine learning algorithms.

3.2 Phase 2 algorithm summary of tests

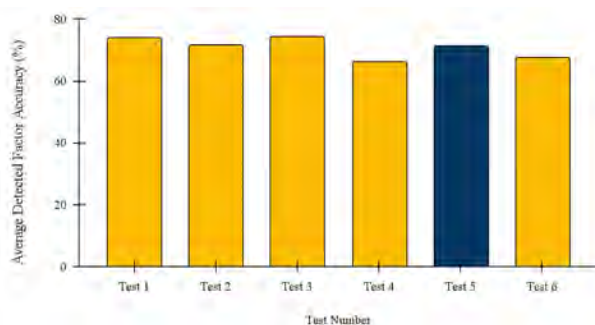


Figure 5. Average Detected Factor Accuracy (%) vs. Test Number

The Phase 2 algorithm summary table shows variance over tests, but a point of maximum accuracy and lowest error variances at Test 5 (highlighted in blue). A total algorithm accuracy rate of 74.3% was calculated from the above values.

100% accuracy was found for all individual Hamming Distance, Greedy, Naïve Brute Force, and Jaccard Index algorithms used throughout all phases. This accuracy rate refers to the baseline working of the algorithms. When a training set of data was used to initially test these algorithms before implementing real-world data, the algorithms produced accurate results matching exactly what was expected of them.

4. Discussion

Considering both algorithms developed in Phase 1 and Phase 2 relatively accurate (in comparison to current algorithms present on drug target identification), two variant propositions to reverting the aberrant case of HCC are founded. First, to induce the 5hmC pathway at positions 33-35. Second, to inhibit the Prp9 spliceosome factor as an acting protein in the spliceosome. Both tactics are predicted to reduce hypermethylation. Due to the fact that neither Phase 1 nor Phase 2 algorithms are predicted to have a 100% accuracy rate, these proposed solutions cannot be guaranteed. In this way, this research serves as more of a basis for novel HCC drug target identification rather than as a cemented solution. Despite this, it provides a tactic of *in silico* drug target identification never previously explored. While methylation and spliceosome factors have been previously studied, it had been work done over

decades, only limited to a taken case study (Xu, et al., 2017). Such efficient case-specific analysis has never been done before. While modern-day drug target computational analyses are only able to narrow down to the top 50-100 leads, this research provides an accurate way to narrow down to 2-5 leads. Such narrowing down saves millions of dollars and years of time spent on drug research.

5. Conclusion

Through this research project, 2 specific algorithms were developed to revert the hypermethylation present in HCC cases. The Phase 1 algorithm focuses on identifying points of hypermethylation in a given sequence to trigger point-reversion utilizing the 5hmC demethylation pathway. The Phase 2 algorithm identifies spliceosome factors responsible for DiR inhibition, a characteristic of hypermethylation, using machine learning. Accuracy rates for the given phases came out to 87.5% and 74.3%, respectively. Algorithms developed in both phases were then applied to the TP53 gene to identify 2 specific propositions to reverting hypermethylation in HCC.

Acknowledgment

Annalisa DiRuscio, Ph.D., MD - Initial research on the role of DiRs.

References

- Chen, Z. X., and Riggs, A. D. (2011). DNA methylation and demethylation in mammals. *J. Biol. Chem.* 286, 18347–18353. doi:10.1074/jbc.R110.205286.
- Dong, E., et al. (2012). Upregulation of TET1 and downregulation of APOBEC3A and APOBEC3C in the parietal cortex of psychotic patients. *Transl. Psychiatry* 2, e159. doi: 10.1038/tp.2012.86
- Dreval, K., et al. (2019). Gene Expression and DNA Methylation Alterations During Non-alcoholic Steatohepatitis-Associated Liver Carcinogenesis. *Frontiers in Genetic.* doi: 10.3389/fgene.2019.00486
- Fan, G., et al. (2018). DNA methylation biomarkers

for hepatocellular carcinoma. *PubMed Central*. doi:10.1186/s12935-018-0629-5

Hlady, R., et al. (2019). Genome-wide discovery and validation of diagnostic DNA methylation-based biomarkers for hepatocellular cancer detection in circulating cell free DN. *PubMed Central*. 9(24), 7239–7250. doi:10.7150/thno.35573

Liu, J., et al. (2016). Abundant DNA 6mA methylation during early embryogenesis of zebrafish and pig. *Nature Communications*, 7(866), 1–7.

Shi, D., et al. (2017). New Insights into 5hmC DNA Modification: Generation, Distribution and Function.

Frontier Genetics. doi: 10.3389/fgene.2017.00100

Wang, Y., et al. (2015). Mechanism of alternative splicing and its regulation (Review). *Biomedical Reports*, 3, 152-158. doi: 10.3892/br.2014.407

Wu, X., & Zhang, Y. (2017). TET-mediated active DNA demethylation: Mechanism, function and beyond. *Nature Reviews Genetics*, 18(9), 517–534.

Xu, Rh., et al. (2017) Circulating tumour DNA methylation markers for diagnosis and prognosis of hepatocellular carcinoma. *Nature Mater*, 16, 1155–1161. doi: 10.1038/nmat4997

Dendrimers and Cancer Therapy: A Review and Analysis

Alexander Xu^{1,2,*}, Emily Reynolds¹

¹Apex Learning Virtual High School, Richardson, TX USA

²Mountain View High School, Mountain View, CA USA

Received July 31, 2022; Revised January 5, 2023; Accepted, January 24, 2023

Abstract

Over the past two decades, dendrimers have emerged as a novel class of nanoparticles for cancer therapy and detection. They come in many varieties and are used to carry molecules like drugs or to act as imaging agents. They are composed of branches radiating out from a core, holding drugs between their branches or at the ends of them. Dendrimers are useful because of their customizability. They can be used for actively delivering drugs to tumors or for detecting cancer. Dendrimers have various traits that make them good for drug delivery, including their stability, water solubility, and uniformity. However, dendrimers also have limitations, displaying toxicity, or noting accumulate enough in the intended areas, and their interactions with the body are little known. Therefore, modifications to dendrimer structures are underway to solve some of the issues that their use presents. Overall, dendrimers show great potential in cancer therapy, but additional work needs to be done to enhance their specificity and effectiveness. They are important in the design of personalized cancer therapies, which we believe are the future of effective cancer treatment.

Keywords: Dendrimers, Cancer Therapy, Nanoparticle, Drug Delivery, Nanotechnology

1. Introduction

Cancer is a disease involving the rapid reproduction of mutated cells, often causing assorted physiological problems by disrupting organ functions. It is considered as the second largest contributor to worldwide mortality. The International Agency for Research on Cancer (IARC), an agency of the World Health Organization (WHO) projected around 29.5 million incidences and 16 million deaths by 2040. (Saluja, et al. 2021) Dendrimers are effective vectors for cancer therapy because they are compatible with the human body and because they are capable of delivering drugs or imaging agents to cancer sites and having a controlled release.

Dendrimers are nanoparticles ranging from 1 to

20 nm in size. The anatomy of dendrimers consists of branches radiating outwards from a core, with surface groups at the periphery. (Li, et al., 2017) Drugs can be stored inside the dendrimer between the branches, or attached to the edges outside the dendrimer. Their ability to be grown from scratch means that the structures of dendrimers are modifiable and more easily customized to suit one's needs. (Ruiz, et al., 2014) Dendrimers, while passively growable, can also be actively modified and rearranged. (Wolinsky and Grinstaff, 2008; Svenson, and Tomalia, 2012; Saluja, et al., 2021; Kumar, et al., 2020)

Dendrimers's solubility, uniformity and biodegradability make them ideal for drug delivery to cancer cells. Dendrimers typically enter cells via endocytosis. In cancer therapy, dendrimers see

* Corresponding Author
alexanderxu1111@gmail.com

Advisor: Emily Reynolds
Emily.Reynolds@edmentum.com

potential as imaging agents and drug delivery vehicles. They can migrate to tumors and allow them to be detected with machinery like X-rays. Additionally, they can be used to deliver drugs to tumors, thus killing them. Normally the release of drugs from the inside is hard to control, but this is solved via photodynamic or photothermal therapy in which certain kinds of light or heat act as triggers to release the drugs. (Wolinsky and Grinstaff, 2008; Svenson, and Tomalia, 2012; Saluja, et al., 2021; Kumar, et al., 2020)

Dendrimers have some issues, most prominently their toxicity, (Ruiz, et al., 2014; Kumar, et al., 2020) and solutions to this are still being worked on. Additionally, dendrimers are more prone to defects on larger scales and are not particularly effective at locating to and accumulating in tumors. (Saluja, et al., 2021; Bugno, et al., 2015) Currently, there are also not enough studies relating to long-term effects of dendrimers in the body, and their interactions with bodily systems are not well-understood. (Lapieciec-Watala and Watala, 2015) This review focuses on the introduction of dendrimers and an analysis of their characteristics that make them good agents for cancer therapy.

2. Dendrimers

Dendrimers are well defined, nano-size, monodisperse, radially symmetric globular architecture having tree-like branching units. Their name is derived from the Greek words dendri-meaning 'tree-like' and meros meaning 'part of' (Kumar, et al., 2021). Dendrimers' important characteristics make them promising in cancer therapy.

2.1 Types of Dendrimers and Dendrimer Anatomy

There are multiple kinds of dendrimers, including polyamidoamine (PAMAMs), Poly (propylene imine) (PPI), Poly-L-lysine (PLL), Core shell (tecto), hybrid dendrimers, and peptide dendrimers (Figure 1). They are small, radially symmetric, and consist of branches sticking out from a central core. (Wolinsky and Grinstaff, 2008) Their general anatomy consists of a core, branches, and an outside surface (Figure 2A).

(Saluja, et al., 2021; Kumar, et al., 2020; Li, et al., 2017; Xu, et al., 2014)

Almost all dendrimers are capable of delivering drugs, and they all have their own up and downsides. Dendrimers deliver a variety of drugs, not restricted to but including cisplatin, doxorubicin, and various photosensitizing drugs. (Gillies and Fréchet, 2005) They can carry drugs inside them, sometimes with the help of hydrogen bonding (Svenson, and Tomalia, 2012) or drugs can be attached to the ends of the branches. (Saluja, et al., 2021) Dendrimer-drug conjugates are dendrimers with drugs on the ends, allowing multiple drug molecules to be attached to the ends of the dendrimer; and unlike drug-encapsulated dendrimers, the release of the drugs are controllable based on the linkages. (Wolinsky and Grinstaff, 2008) For cancer treatment, the most commonly used dendrimers are PAMAMs, PPI, PLL and Peptide drimers. (Saluja, et al., 2021) PAMAMs with ammonia/ethylene diamine as a core are a versatile tool due to their monodispersity. They are good at probing for cancer, as they immobilize DNA aptamers. (Li, et al., 2017) AMAM-organosilicon dendrimers have PAMAMs on the inside and organosilicon peripherals. (Kumar, et al., 2020) PPIs are made of poly- alkyl amines inside and tertiary tris-propylene amines on the outside. They are hypothesized to be more effective at holding lipophilic drugs. Peptide dendrimers consist of a peptidyl core and peptide chain peripherals, and can be used for diagnosis as well as vaccine/gene delivery. Tables 1 and 2 show examples of dendrimer-based nanocarriers used for nucleic acid and drugs. (Palmerston Mendes L., et al., 2017)

2.2 Dendrimer growth and modification

Dendrimers are stable, capable of self-assembly, and polyvalent. (Kumar, et al., 2020) When grown, each activation and condensation reaction represents a step or a generation in which the number of surface groups doubles. (Bugno, et al., 2015) Figure 2B showed linear increase in diameter and exponential growth of the number of surface groups. (Xu, et al., 2014)

There are two main ways that dendrimers are grown: Either the dendrimer can be grown outwards

from the core (the divergent method), or the dendrimer can be grown inwards from the peripheral regions (the convergent method). (Wolinsky and Grinstaff, 2008; Saluja, et al., 2021) The larger a dendrimer is, the more spherical it becomes. (Kumar, et al., 2020) The amount of methods of growing dendrimers has led to over a hundred families and a thousand different surface modifications. () Two other emergent modifications include “Lego chemistry,” which seemingly involves the use of pre-made cores and branches to create a dendrimer, as well as “click chemistry” which allows for the synthesis and isolation of dendrimers. (Svenson and Tomalia, 2012)

2.3 Dendrimer properties and development

Some key properties of dendrimers as nano-carriers of cancer drugs are being identified. Dendrimers are known for having a very controllable structure. (Li, et. al., 2017) Their “void spaces” between the branches are valuable because they allow the dendrimer to act as a transport vessel for drugs that are too insoluble to dissolve in the body by themselves. This is helped by the fact that the dendrimer itself is soluble. Additionally, they also help protect the drugs from destruction before they reach their targets. The ability for drugs to be stored within dendrimers allows for the development of controlled release methods that allow for more targeting precision. Dendrimers have a defined, uniform structure, which allows them to be reproduced and changed more easily. (Kumar, et al., 2020) They additionally are very small, being nanomaterials, making them better at crossing the cell membrane and reducing their clearance from the body. Dendrimers have long blood plasma retention time and are able to circulate for hours before being excreted, greatly influencing drug pharmacokinetics. These properties of dendrimers make them ideal candidates for cancer drug delivery. (Saluja, et al., 2021; Kumar, et al., 2020)

Recently, dendrimers have progressed in development, and there are now some improvements being made to them. Modifying the charge of peripheral groups can help the dendrimer enter the body as well as help them spread throughout it.

However, they cannot carry large amounts of drugs inside of them. Janus dendrimers are dendrimers with distinct groups on them that are different in terms of chemicals and structure, which can allow for the addition of multiple kinds of surface groups on one dendrimer or the higher loading of hydrophobic drugs, but in general drugs still have to be conjugated to be used inside of them; a similar problem that occurs with normal dendrimers. The merging of dendrimers with other nanocarriers has been proposed, in which dendrimers are put inside an even larger nanoparticle, leading to quicker circulation and better tumor accumulation. (Bugno, et al., 2015)

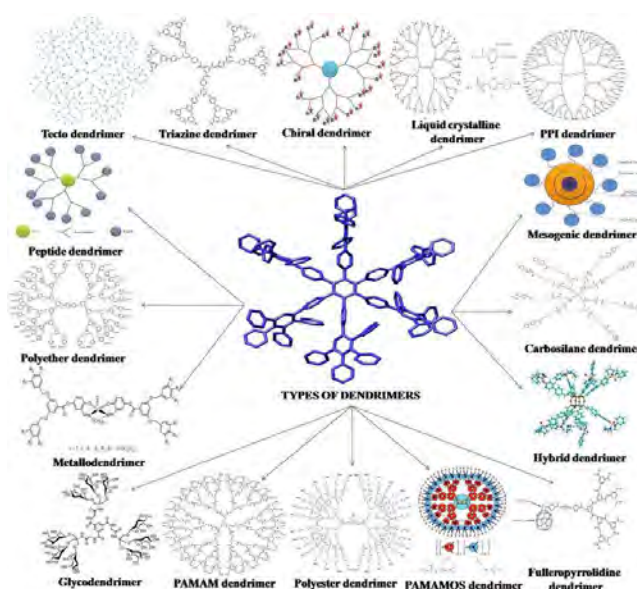


Figure 1. Major types of dendrimers. (Kesharwani, et al., 2014)

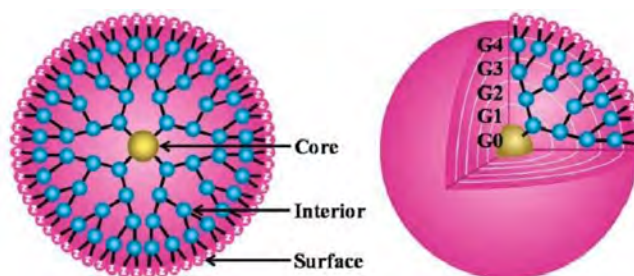


Figure 2. (A) Schematic presentation of dendrimers as a nanoscaffold with a core, interior branches and an outside surface (B) The growth of a dendrimer from its core. Abbreviations: G = generation; Z = surface group for host-guest interactions and functionalization. (Xu, et al., 2014)

3. Cancer Treatments and Detection

3.1 Dendrimers in cancer therapy

Dendrimers have been investigated extensively in the medical field, and cancer treatment is one of the areas where they have been most used. They are highly branched polymers with easily modifiable surfaces, which permits conjugation with drugs and DNA/RNA used for cancer therapy. (Palmerston Mendes L., et al., 2017) The drugs can be stored in the cavity of the dendrimers or conjugated to their functional groups at their surface. Nucleic acids usually form complexes with the positively charged surface of most cationic dendrimers (Table 1). On the other hand, folate and antibodies are often added as moieties on the dendrimer’s exterior (Table 2).

Table 1. Examples of dendrimer-based nanocarriers for nucleic acids

Polymer	Generation	Payload	Application
PAMAM	G1	replicon mRNA	Vaccine
	G4	siBCL-2	Ovarian cancer
	G4	IFN-b	Malignant glioma
PPI	G2	pCN-Luci	LHRH positive cancer
	G3	siBCL-2	Lung cancer
	G5	siBCL-2	LHRH positive cancer
PLL	G3-G6	p-Luci	Gluconeogenesis
	G6	siGAPDH siOCT1	Gluconeogenesis
	G6	pCMV-Luc	Gluconeogenesis

Table 2. Examples of dendrimer-based nanocarriers for drugs.

Polymer	Generation	Payload	Application
PAMAM	G4	Doxorubicin	Lung metastasis
	G4	Paclitaxel	Breast cancer
	G4	Docetaxel	Breast cancer
PPI	G3-5	Melphalan	Breast cancer
	G4.5	Paclitaxel	Ovarian cancer
PLL	G6	Doxorubicin	Melanoma
	G6	Doxorubicin	Rectum cancer

3.2 Dendrimers in cancer drug delivery

In cancer therapy, dendrimers are intended to be used in order to deliver drugs used for different types of cancer. These drugs are either conjugated to the molecular structure of the dendrimer itself, or enclosed within the spaces between the branches of the dendrimer. (Wolinsky and Grinstaff, 2008) Targeted delivery of dendrimers to tumor cells prevents harmful side effects like toxicity.

There are two main kinds of targeting, passive or active. Passive targeting exploits the often inefficient or deformed properties of tumors, such as changed permeability of tumor blood vessels, which leads to the dendrimers naturally being absorbed and retained inside the tumors. Alternatively, active targeting involves the use of conjugation of targeting moieties to the dendrimer so that they can bind to the cell receptors of the cancer cells to enter. (Wolinsky and Grinstaff, 2008; Ruiz, et al, 2014) Scientists have studied the results of conjugating monoclonal antibodies, common cancer drugs, to dendrimers to target tumor cells that express certain antigens. For example, Human growth factor receptor-2 (Her-2) is often overexpressed in breast and ovarian malignancies. Shukla et al. showed that when an anti-HER2 was conjugated to G5-PAMAM and evaluated for binding affinities and internalization, it was found to specifically bind to HER2-expressing cells both in vitro and in vivo (Shukla, et al., 2006) It has been proposed that the shape of dendrimers gives them the ability to enter a cell via the cell membrane, but the more conventional method of entering a cell is through endocytosis. (Kumar, et al., 2020)

3.3 Dendrimers in photodynamic and photothermal cancer therapy

A major potential application of dendrimers is in photodynamic therapy, in which a photosensitizing agent is activated via light exposure and becomes toxic enough to cause cell death in tumors by generating reactive oxygen (Gillies and Fréchet, 2005; Ruiz, et al., 2014). Dendrimers have the potential to increase the efficiency of the photosensitizing agent’s delivery, with the ideal dendrimer producing lots of singlet oxygen, no

toxicity when not exposed to light, high solubility, tumor targeting, and the ability to absorb light. (Klajnert, et al., 2012) Dendrimer- micelle complexes are especially effective, as they accumulate in areas prone to tumors and specifically target the vulnerable organelles of the cancer cells. (Wolinsky and Grinstaff, 2008)

Photothermal therapy operates in a similar way, except the dendrimers contain metal nanoparticles that heat up to lethal temperatures when exposed to light, killing tumor cells. (Wolinsky and Grinstaff, 2008) These strategies are very useful because they allow for complete control over drug release. (Ruiz et al., 2014) Boron neutron capture therapy is another option that uses boron-10's neutron capture reactions to cause localized cell death. (Kumar, et al., 2020)

3.4 Dendrimers in cancer detection

Dendrimers do not only have potential in cancer treatment, they may also be useful in cancer detection. By tethering fluorescent molecules to dendrimers, they can make their way to cancer cells, and their path can be tracked via fluorescence spectroscopy. Dendrimers with high molecular weights may make good MRI contrast agents. Their water solubility and stability also makes them good X-ray contrast agents, especially when iodine and barium are incorporate in CT and X-ray imaging, although the materials used could be toxic. (Kumar, et al., 2020)

3.5 Advantages of dendrimer drug delivery

Together with surgery and radiation, chemotherapy is one of the most commonly used cancer treatment options in the clinic. Chemotherapy drugs are frequently used for suppressing cancer cell growth, metastasis and disease progression. However, these drugs kill not only proliferating cancer cells but also normal cells. Other drawbacks include low solubility and bioavailability, short circulation time and nonselective biodistribution, which lead to unfavorable side effects. (Li, et al., 2017)

Compared to traditional drug delivery systems, dendrimers have shown greater potential in improving drug bioavailability, targeting tumors with

specificity, prolonging drug circulation time and controlling drug release. Another major benefit to using dendrimers is the amount of control that they can be created with; as they have a modular and easily-modified structure (Ruiz, et al., 2014), which gives them lots of potential in drug delivery and cancer therapy. They are manipulable by changing the end groups and increasing generation, as well as modifying the mass of the components like the branches (Wolinsky and Grinstaff, 2008; Svenson, and Tomalia, 2012) and the potential diversity of forms means that it is easier to experiment with them until a desired result is reached. Additionally, they are uniform and have a well-defined structure; and as they are hydrophilic, they are also water- soluble. (Kumar, et al., 2020) They can be easily rebuilt. (Li, et al., 2017) Additionally, as multiple drugs can be placed on a dendrimer, they have a greater effect over other nanoparticles. Because drugs can be stored inside dendrimers or even in the core, they are shielded from the environment. (Klajnert, et al., 2012)

3.6 Disadvantages of dendrimer drug delivery and their solutions

Dendrimers have some drawbacks. First of all, there are some issues with the most commonly-used form of synthesis, divergent synthesis, which is prone to defects on larger scales. Larger dendrimers are more likely to have structural issues. (Bugno, et al., 2015) High generation dendrimers often have structural problems, although methods that fix this have been developed. (Saluja, et al., 2021)

Dendrimers can be quite toxic, especially large, catatonic ones due to their stability. (Ruiz, et al., 2014; Kumar, et al., 2020), and therefore harmful in large amounts, with surface groups often having to be modified to reduce toxicity and liver accumulation. (Gillies and Fréchet, 2005; Labieiec-Watala and Watala, 2015) They form holes in the lipid bilayer of cells; so the end groups should be capped to avoid injury to them. (Wolinsky and Grinstaff, 2008) Dendrimer interaction with blood is largely unknown, and few studies looking at the long-term effects of dendrimers have been carried out.

Drug-encapsulated dendrimers are difficult to

control, either releasing the drugs over several hours (Wolinsky and Grinstaff, 2008), releasing the drugs immediately, or requiring the dendrimer to face “harsh conditions” (Gillies and Fréchet, 2005), so they are most effectively used by injecting them directly into the tumor. Attempted solutions to the problem of controllability include the use of stabilizing polyethylene oxide chains or the hybridization of polyethylene oxide with dendrimers. Additionally, dendrimers do not accumulate in tumors enough, oftentimes with only 10% of the dose making it to its intended location. (Bugno, et al., 2015)

Photodynamic therapy also has issues, as photosensitizers can be toxic to important tissues and don't always target tumors specifically. α -lipoic acid (ALA) could be used to improve selectivity, but ALA moieties have been shown to increase the amount of toxic material generated inside tumors. The addition of micelles seemingly increases the amount of dendrimers accumulating in tumors, although they are ineffective at carrying drugs. Additionally, the opacity of skin makes sufficient light exposure difficult. The light penetration issue is being solved by lasers. (Gillies and Fréchet, 2005)

4. Conclusions

Dendrimers are continuing to be developed, built upon, and refined, with attempted improvements and new variations often appearing. The amount of control that researchers have over dendrimer structure, and thus function and usefulness, is a defining feature of dendrimers. (Wolinsky and Grinstaff, 2008) In recent years, there have been water soluble and biocompatible dendrimers created as well, which makes them far less harmful to healthy tissues and thus making it safer to use them to treat cancer. Their selectivity - only delivering drugs to cancerous tissues - would also give them an advantage over other cancer drugs. (Saluja, et al., 2021) They have so much potential in anti-cancer therapy because of their uniform size, distribution, and composition. (Ruiz et al., 2014) Photodynamic and photothermal therapy, as well as simply using dendrimers to deliver drugs, are all promising areas of research within the subject of using dendrimers in

cancer therapy. Continued research in the area will bring macromolecules with increasing specificity and efficacy towards the diagnosis and treatment of cancer in the clinic.

In my opinion, the first major area to improve is specificity; as anything resembling a one-size-fits-all treatment is likely impossible due to the variety of cancers, variation of cancer microenvironment and difference in patients' response to different therapeutics. The most appropriate approach should be chosen based on individual cases and characteristics of each tumor. The multivalency of dendrimers is important in the design of personalized therapies. A customized dendrimer structure can potentially be tuned simultaneously for desired biocompatibility, bioavailability, drug circulation time, and targeted delivery of therapeutics to tumors. Efficacy could perhaps be improved with certain surface-level modifications, by finding a way to increase drug load, or by incorporating dendrimers into a more effective structure. In summary, there is no one technology that solves all issues. Instead, a patient-specific therapeutic agent would be required to achieve effective cancer treatment in years to come.

Acknowledgement

I would like to acknowledge Miss Emily Reynolds, who is my teacher and advisor for this independent research project. Her insights and support made the article possible.

References

- Bugno, J., Hsu, H.J., & Hong, S. (2015) Tweaking Dendrimers and Dendritic Nanoparticles for Controlled Nano-bio Interactions: Potential Nanocarriers for Improved Cancer Targeting. *Journal of Drug Targeting*, 23(7-8), 642-650.
- Gillies, E.R., & Fréchet, J.M.J. (2005) Dendrimers and dendritic polymers in drug delivery. *Drug Discovery Today*, 10 (1), 35-43.
- Kesharwani, P., Jain, N. K., & Jain, K. (2014) Dendrimer as nanocarrier for drug delivery. *Progress in Polymer Science*, 39(2), 268-307.

- Klajnert, B., Bryszewska, M., & Rozanek, M. (2012) Dendrimers in Photodynamic Therapy. *Current Medical Chemistry*, 19(29), 4903-4912.
- Kumar, V., et. al. (2020) Diagnostic and therapeutic applications of smart nanocomposite dendrimers. *Frontiers in Bioscience-Landmark*, 26(3), 518-536.
- Kumar, K., Singh, A., & Gupta, A. (2021) A role of dendrimers in drug delivery for cancer therapy. *Int. J. Indig. Herbs Drugs*, 1(6), 9–16.
- Labieiec-Watala, M., & Watala, C. (2015) Dendrimers: Destined for Success or Doomed to Fail? Plain and Modified PAMAM Dendrimers in the Context of Biomedical Applications. *Journal of Pharmaceutical Sciences*, 104(1), 2-14.
- Li, Z. et. al. (2017) Cancer drug delivery in the nano era: An overview and perspectives (review). *Oncology Reports*, 38(2), 611-624.
- Palmerston Mendes, L., Pan, J., & Torchilin, V.P. (2017) Dendrimers as Nanocarriers for Nucleic Acid and Drug Delivery in Cancer Therapy *Molecules*, 22(9),1401-1420.
- Ruiz, M.E., Gantner, M.E., & Talevi, A. (2014) Applications of nanosystems to anticancer drug therapy (Part II. Dendrimers, micelles, lipid-based nanosystems) *Recent Patents on Anticancer Drug Discovery*, 9(1), 99-128.
- Saluja, V., et. al. (2021) Dendrimers based cancer nanotheranostics: An overview. *International Journal of Pharmaceutics*, 600(2021), 120485.
- Shukla, R., et al. (2006) HER2 specific tumor targeting with dendrimer conjugated anti-HER2 mAb, *Bioconjug. Chem.* 17(1) 1109–1115
- Svenson, S. & Tomalia, D.A. (2012) Dendrimers in biomedical applications—reflections on the field. *Advanced Drug Delivery Reviews*, 64(2012), 102-115.
- Wolinsky, J.B., & Grinstaff, M. (2008) Therapeutic and diagnostic applications of dendrimers for cancer treatment. *Advanced Drug Delivery Reviews*, 60(9), 1037–1055.
- Xu, L., Zhang, H., & Wu, Y. (2014) Dendrimer advances for the central nervous system delivery of therapeutics. *ACS Chemical Neuroscience*, 5(1), 2-13.

The Unknown Death Toll of COVID-19: How The Pandemic Killed One Million Americans Long Before 2022

John Dong^{1*}

¹Campolindo High School, Moraga, CA USA

Received June 29, 2022; Revised January 25, 2023; Accepted, February 7, 2023

Abstract

In this paper, I intend to convince the reader that the actual number of deaths caused both directly and indirectly by the COVID-19 pandemic in the United States of America has been much more than officially reported. According to an excess death regression model which accounts for circumstantial and unreported COVID deaths, there are approximately 130,000 additional pandemic-related deaths in 2020 and 140,000 in 2021. Adding the officially reported 831,000 COVID deaths in these two years, the pandemic may have caused over 1 million deaths much sooner than the officially reported date, May 24th, 2022.

Keywords: Bioinformatics; Computational Biomodeling; Coronavirus; Death Toll; Excess Deaths

1. Introduction

As of April 5th, 2022, approximately 981,000 Americans have passed away from COVID-19 (Elflein, 2022). Such a statistic is already dismal enough, but these deaths only include deaths that are directly listed as from COVID-19. The true number of deaths is in actuality far higher for a multitude of reasons. Faulty or inconsistent reporting has, of course, resulted in statistical snafus. While most sources give off relatively accurate and reliable information that is recorded to the best of the government's ability, they still have significant shortcomings. Much of the time, information is simply not up to date. According to the Center for Disease Control and Prevention, it can take up to 60 days for a death to be reported and logged in government records (CDC, 2020). A study conducted by the Cable News Network claims that COVID positive and mortality results are often sent by fax, email or even snail mail, meaning that staff and volunteers can take several days to log the

information (McPhilips, 2022). Human random error was also present, as staff logging two people with the same name as one patient or incorrectly logging dates is not unknown. Sometimes, COVID reports are simply lost to time and space due to system malfunctions or inexplicable gaps in recording (Banco, 2021). Then there is also the question of what defines a COVID death. Deaths from pandemic-related complications, such as those with underlying health conditions exacerbated by COVID, are often not reported as COVID deaths and therefore not included in the official death count. It seems reasonable to me to count such deaths as COVID deaths, but this sentiment sees little praxis officially. The cause of death is often hard to directly ascertain in many cases, and as a result deaths from coronavirus complications are often not listed as from COVID. Many COVID deaths have also been from other underlying conditions, not COVID itself.

Additionally, direct mortality caused by COVID infection is only but a facet of the death toll in reality. Not often considered is the economic aspect: the

* Corresponding Author
theclockworkbishop123@gmail.com

Advisor: Dr. Ganesh Mani
ganeshm@andrew.cmu.edu

effects of the pandemic on the economy have emphatically led to a higher death rate amongst the American population. In 2021, millions of Americans lived paycheck to paycheck. The federal minimum wage has not been adjusted for the skyrocketing cost of living (Desilver, 2017). This means a large number of Americans were living on the brink even before COVID. However, the pandemic has exacerbated the already dismal homelessness crisis (11). Along with the layoffs and downsizing, many workers who tested positive for the virus were left without a source of income. Megan Ranney, M.D., summarizes that “...for much of America, living with long Covid is enough to put folks over the edge financially, with very limited safety nets.” Many stories have surfaced on the internet of people who have lost their jobs, homes and livelihoods due to the pandemic (12). Consequently, homeless deaths have increased directly (13). Furthermore, a 2009 Harvard Medical School study found that 45,000 Americans die every year as a direct consequence of lacking health insurance coverage. A December 2019 poll claims that 25% of Americans have had delayed medical treatment because of the cost (Sainato, 2020). 100,000 American diabetics die each year from the cost of insulin, and insulin costs soared as a direct result of the pandemic (Marquardt, 2022). With the waves of layoffs and overwhelmed hospitals in the pandemic’s wake, these numbers have increased, resulting in a larger death toll for those who are not able or cannot afford to access healthcare (Chokshi and Murthy, 2022). The aforementioned factors mean that the number of deaths in truth is most certainly higher than claimed.

These are all examples of mortality from “COVID related” causes, and not the coronavirus itself. Such numbers are typically not reported in official death counts, as Figure 1 will attest to. Thanks to Figure 1’s highly accessible and readable nature, it is easy to see why so many people would be misled about the true death count of the pandemic. This underreporting of COVID mortality reveals a broader issue: access to accurate and authentic information. Despite it being easy to obtain death statistics from various online resources, it is not straightforward for the general public to estimate the actual numbers of people who die of COVID. Hyper-individualist movements like

the anti-maskers and anti-vaxxers have emerged as a result of distrust in mainstream media sources and inaccurate reporting. They may be the face of this growing insecurity about misinformation, but of course they do not necessarily represent the American population in its entirety. Still, roughly half or more Americans do get their information about COVID from online or social media sources (4). And when COVID death counts are so drastically underreported as mentioned above, there is little doubt that most Americans will drastically underestimate the total number of pandemic related deaths.

Therefore, the subject of this paper is to convince the reader of the true size and scale of the pandemic’s human toll in America. By analyzing morality data from the years 2020 and 2021, an estimate can be made about how many pandemic related deaths have not been reported or commonly considered in official death counts, with causes ranging from unreported COVID infection to the other effects of the pandemic.

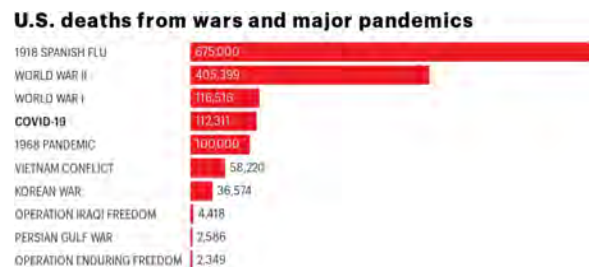


Figure 1. Easily digestible infographics like such only show the recorded number of confirmed COVID deaths and completely ignore the excess mortality resulting from the other effects of the pandemic (Lambert, 2020).

2. Materials and Methods

The main crux of the research conducted was analyzing excess mortality data for the entirety of the United States of America. Excess mortality is defined as the number of deaths that exceed the expected or “normal” number of deaths. The expected number of deaths is usually calculated through analyzing previous mortality data and finding a general estimate or trend. Using excess death data can mitigate the

effects of poor recording or faulty labeling, since death certificates will be reported regardless of the cause. Excess deaths are defined as the deaths outside of predicted numbers, which would then include pandemic-related deaths. Political tensions, random error and system malfunctions can also cause significant delays of up to sixty days for COVID case logging. Complications resulting from COVID and death from lack of adequate healthcare as a result of the pandemic are not recorded in official confirmed COVID death counts. This means that confirmed COVID deaths are likely to experience statistically significant effects due to logistic failures, and cannot be relied upon as a true measure of the pandemic's net mortality. Therefore, comparing excess mortality and expected mortality rates is the most accurate and objective form of analysis, as it factors in both underreported COVID deaths as well as pandemic-related deaths with minimal error.

Official CDC government records were chosen, as they are the most easily accessible and generally accurate (CDC, 2017). All recorded and confirmed COVID deaths will be excluded from consideration in the scope of the research. This data is essentially irrelevant, as confirmed cases are already counted in most conventional estimates. I am only trying to find the deaths that are not commonly considered so as to not be redundant. I have chosen only 2020 and 2021 to be the scope of my study, as data from 2022 appears too recent and cannot be determined to be as accurately reported. As has already been mentioned, significant delays in reporting data is not uncommon. Thus, the data is established to be the number of excess deaths excluding confirmed COVID cases, with the desired numbers being the pandemic-related deaths.

A simple regression model can be used to compare excess mortality with predicted mortality rates. Regression is essentially a model or function that describes the relationship between an independent and dependent value quantitatively, and can be used to predict future occurrences. In other words, it takes numerous pieces of data and finds a general trend, which can then be used to predict future occurrences. Through the Python numpy-based Seikit-Learn modification, a regression curve of best

fit can be applied to a set of data, and then compared with another set of data. In my experiment, I will first use a sinusoidal regression model to fit the total 2019 pre-pandemic death counts recorded per week and establish a general trend curve. The 2019 pre-pandemic death data would be a reasonable prediction for 2020 and 2021, since population growth was negligible. Official CDC data states that the trend of mortality rates holds constant or at least similar for the past decade, so the regression model used the mortality data for the pre-pandemic year 2019 as a basis for training and recognition. Population growth in 2019 was roughly 0.6%, and population growth rates in America have been declining, to the point of almost flatlining in the last decade. Since in 2020 and 2021 the population growth was less than 0.1%, this new growth have been assumed to be not statistically significant to the model. This will be the benchmark for which I will then use the regression trend curve and superimpose it over total non-COVID mortality data from 2020 and 2021, and then contrast. The excess amount will obtain the total number of pandemic-related deaths that were not officially listed as from COVID-19. The CDC has also provided rough estimates of predicted mortality rates, based on 6 years of previous data and adjusted to population growth. This will be another form of analysis that can be used to determine the rough number of pandemic-related deaths.

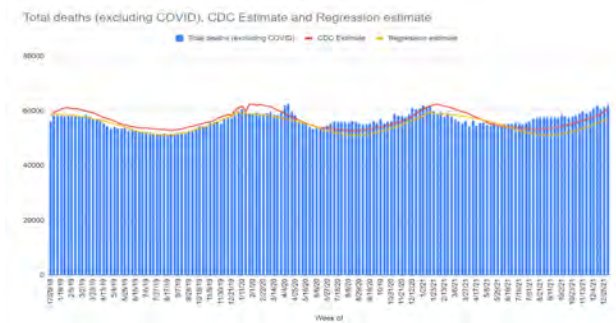


Figure 2. A bar chart of recorded total death count excluding those OFFICIALLY listed as COVID compared with the CDC estimate curve and regression estimate. Estimated excess deaths are denoted by the total death counts that exceed the estimates.

3. Results

January 20th was when the first lab-tested coronavirus case was confirmed in the US, and roughly a month later February 29th the first death. The data shown in Figure 2 is a representation of all recorded excess deaths from 2019 to 2021, excluding confirmed COVID-19 counts. My regression estimate, represented by the yellow line, has been curve fitted with only data from 2019, since that year is assumed to be COVID-free. (Some sources have claimed that the coronavirus was present in America in 2019. However, these cases are generally few and far between, and assumed to be statistically negligible.) The CDC estimate, in red, is based on six years of previous data. It consistently overestimates the actual number of deaths, with an average percent of error of 2.91%. In contrast, the regression estimate used in this paper only has a 0.223% percent error, suggesting it can be used to accurately predict the deaths count trend.

4. Discussion

When comparing recorded non-COVID death counts with expected total deaths counts from 2020 to 2021, there is a clear difference. A comparison of the regression model's expected death trend line and the recorded number of total non-COVID deaths for 2020 yields approximately 130,000 excess deaths in 2020, and 140,000 in 2021, as seen in Figure Three. The percent error of 0.223% can be generally assumed to be negligible. The CDC estimate yields approximately 40,000 deaths above the expected number in 2020. However, one must keep in mind that the CDC estimate was consistently much higher than the actual number of deaths for 2019, with a 2.91% error. Therefore, the CDC estimate, when weighted with respect to the error, shows up to 129,000 pandemic-related excess deaths in 2020 and up to 132,000 in 2021.

Viewing the chart, the peak number of overall deaths in 2020 and 2021 is clearly greater than the peak number from 2019, despite population growth being minimal. Although the CDC model took population growth into account, my regression model did not, on the assumption that it would be minimal.

This likely is the reason for my regression model's lower death count. I also did not take into account other confounding variables that may impact death count, such as volume of traffic, suicide rate, crime rate, and other such events that may have been affected by COVID lockdown procedures and such. Therefore, the actual number of pandemic-related deaths will likely be even higher than my estimates.

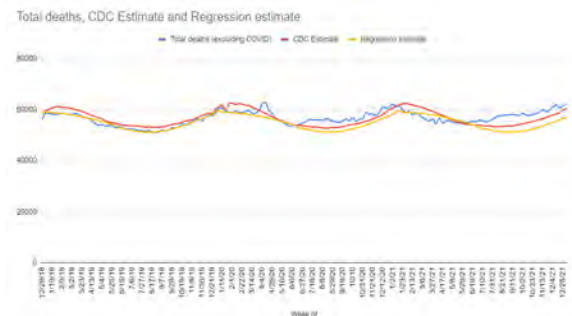


Figure 3. A line chart showing total deaths excluding those officially designated as COVID compared with the curves for the CDC estimate and the regression estimate. Estimated excess deaths are denoted by the total death counts that exceed the estimates. It is clear that in 2020 and 2021 the number of excess pandemic-related deaths, combined with deaths officially labeled as COVID, has far exceeded one million before May 24th, 2022.

5. Conclusion

In 2019, approximately 2.9 million Americans died in total. In 2020, approximately 3.3 million Americans died in 2020. In 2021, there were even more. Roughly 385,000 COVID-confirmed deaths were reported in 2020, and according to my estimates 130,000 other pandemic related deaths occurred. In 2021, there were 446,000 US deaths attributed directly to COVID-19, while my estimates record an additional 130-140,000 pandemic-related deaths. As of May 24th, 2022, the mainstream conclusion is that over one million Americans have officially been reported to have died of COVID-19.

Regrettably, the modeling method employed by said media outlets did not factor in major factors that could have impacted excess death count. For example, vehicular accidents and suicides would have both been impacted by the lockdown policies.

Given the logistical restraints, this data was impossible to implement and represent. Of course, these conclusions are based on the assumption that the government data is reasonably accurate. Such a method may not be employed in regions where government record-keeping is obviously and indubitably neglected.

If these pandemic-related deaths are considered, then well over one million Americans have died from the pandemic. As they say, “the death of one is a tragedy, but a million is a statistic.” It becomes infinitely more heartbreaking when one realizes that one million is composed of one million “ones.” Families have been bereaved, communities torn apart, and the image of a “shining city upon a hill” irreparably shattered.

However, the fact that the pandemic-related deaths are not commonly depicted in official death counts available to the public is of considerable concern. If these effects of the pandemic are ignored or overlooked by the general public, then there can be no possibility for solutions. The pandemic-related deaths did not suddenly materialize. They are the result of long-standing issues like a weak public health system, political tensions, and growing inequalities (Ortiz, 2022). Many of these pandemic-related deaths were preventable and this has significant policy implications for the future.

The accurate tabulation of mortality and related information is almost just as important as medical research itself, and this is why the knowledge of pandemic-related deaths must be widely disseminated. Only with an informed public can there be change. If the information is faulty, incomplete or inconsiderate, then the decisions made from that information will be so too. The American public must realize that death and infection counts are not the only effects of this pandemic, and clamor for positive changes in the future.

Acknowledgements

My sincere gratitude to Dr. Ganesh Mani for the mentorship. His support and advice are greatly appreciated.

References

- Banco, E. (2021, August 15). Inside America’s Covid-reporting breakdown. *The Politico*. <https://www.politico.com/news/2021/08/15/inside-americas-covid-data-gap-502565>
- Center for Disease Control and Prevention. (2017, October 11). Information Quality Support. Center for Disease Control and Prevention. <https://www.cdc.gov/os/quality/support/info-qual.htm>
- Center for Disease Control and Prevention. (2020, June 3). Excess Deaths Associated with COVID-19. Center for Disease Control and Prevention. https://www.cdc.gov/nchs/nvss/vsrr/covid19/excess_deaths.htm
- Chokshi, D., and Murthy, V. (2022, July 2). COVID-19 has derailed health care for many Americans, causing countless deaths. *USAToday*. <https://www.usatoday.com/story/money/economy/2021/09/22/cost-of-living-63-of-urban-americans-live-paycheck-to-paycheck/118830788/>
- City News Service. (2022, April 24). Study: Homeless Deaths Spiked During Pandemic, but Drug Overdoses Remain Top Killer. *NBC News*. <https://www.nbclosangeles.com/news/local/study-homeless-deaths-spiked-during-pandemic-but-drug-overdoses-remain-top-killer/2877200/>
- Desilver, D. (2017, July 1). 5 facts about the minimum wage. *Pew Research Center*. 5 facts about the minimum wage | *Pew Research Center*
- Elflein, J. (2022, May 16). Total number of US COVID-19 cases and deaths. *Statista*. <https://www.statista.com/statistics/1101932/coronavirus-us-covid19-cases-and-deaths-number-us-americans/>
- Lambert, L. (2020, May 15). U.S. deaths from wars and major pandemics. *Fortune*. [nolvi-u-s-deaths-from-wars-and-major-pandemics-1.png \(1024x545\) \(fortune.com\)](https://www.fortune.com/2020/05/15/nolvi-u-s-deaths-from-wars-and-major-pandemics-1.png)
- Marquardt, A. (2022, January 31). There’s a silent epidemic killing over 100,000 Americans a year, and the government is allowing the cure to get more and more expensive. *Fortune*. <https://fortune.com/2022/01/31/how-expensive-is-insulin-diabetes-silent-epidemic-covid-coronavirus/>

Masson, G. (2022, March 1). 43% of Americans have had COVID-19, CDC estimates. *Becker's Hospital Review*.

<https://www.beckershospitalreview.com/public-health/43-of-americans-have-had-covid-19-cdc-estimates.html>

McPhillips, D. (2022, February 21). The US still isn't getting Covid-19 data right. *Cable News Network*.

<https://www.cnn.com/2022/02/10/health/covid-data-problems/index.html>

Miller, J. (2021, January 14). USC study: COVID-19 reduced U.S. life expectancy, especially among Black and Latino populations. *University of Southern California*.

<https://gero.usc.edu/2021/01/14/covid-19-reduced-life-expectancy/#:~:text=and%20Latino%20populations-,USC%20study%3A%20COVID%2D19%20reduced%20U.S.%20life%20expectancy%2C%20especially,by%20more%20than%20a%20year>

Mitchell, A., and Liedke, J. (2021, August 24). About four-in-ten Americans say social media is an important way of following COVID-19 vaccine news. *Pew Research Center*.

<https://www.pewresearch.org/fact-tank/2021/08/24/about-four-in-ten-americans-say-social-media-is-an-important-way-of-following-covid-19-vaccine-news/>

out-four-in-ten-americans-say-social-media-is-an-important-way-of-following-covid-19-vaccine-news/

Moore, S. (2021, Sep 28). History of COVID-19. *News Medical Life Sciences*. (History of COVID-19 (news-medical.net))

Ortiz, J. L. (2022, May 15). 'A terrible price': As US mourns 1 million COVID-19 deaths, preventable losses among unvaccinated a major regret.

USAToday.

<https://www.usatoday.com/in-depth/news/nation/2022/05/15/us-reaches-1-million-covid-deaths/7043778001/>

Sainato, M. (2020, January 7). The Americans dying because they can't afford medical care. *The Guardian*.

<https://www.theguardian.com/us-news/2020/jan/07/americans-healthcare-medical-costs>

Yuko, E. (2022, February 25). First They Got Long Covid. Then, It Made Them Homeless. *Rolling Stone*.

<https://www.rollingstone.com/culture/culture-features/long-covid-homeless-chronic-illness-gig-economy-1312460/>

The Secret World of Crows (*Corvus*): Biology, Behavior, & Cognitive Abilities

Ingrid Lam¹, Marina Seidl^{1*}, Linnea Anderson¹, Pritam Khalsa¹, Leah Jones Neuray¹

¹Northwest Academy, Portland, OR USA

Received June 8, 2022; Revised February 2, 2023; Accepted, February 25, 2023

Abstract

Crows are an enigma to the world of biology—still under-researched and the extent of their cognitive abilities remains unknown. Throughout millennia, crows have influenced human cultures, including artwork, literature, poetry, and language. From reverence to repulsion, human perceptions of crows have evolved throughout history. As a recurring avian symbol in human society, our goal was to summarize recent and historical research on the complex and intricate minds of crows and ravens, while creating artwork in response to our work. In addition, we wanted to inform the public of crows' ecological importance and presence in urban environments. This manuscript synthesizes recent scientific discoveries of crows and ravens and includes our original artwork as figures. Exploring crow behavior and biology, this review paper is organized into five thematic sections characterizing the life of a crow—neuroscience, communication, learning, urban habitat, and interactions with humans.

Keywords: Crow, Human, Behavior, Culture, Neuroscience, Consciousness, Communication

1. Introduction

With a high neural density and well-developed nidopallium, crows possess complex brain structures and intricate neuronal organization that necessitate a more nuanced understanding of avian consciousness. Crow communication is elaborate. From caws to chortles, they produce a variety of sounds used to orally convey valuable information about possible predators, as well as utilizing body movements to inform each other about their surroundings. Patterns of communication are affected by geographic region, social learning, and family dynamics. Crows react to potential threats in their environment by combining the use of detailed spatial memory, tool-making abilities, and caching behavior. Cities offer a plentiful habitat for crows to feed, nest, and avoid predators. Crows often use common city features to their advantage—whether using cars to crack nuts or roadways to drop prey, they are well-adapted to live

within cities. Although crows thrive in an urban matrix, human city residents often view these birds as a nuisance. The objective of this manuscript is to further the audience's understanding of corvid behavior and culture in a way that attempts to reduce negative connotations towards crows while giving readers a glance into the complex workings of these intelligent creatures. This paper discusses novel discoveries in the field of crow behavioral ecology, with the goal of helping the reader become more knowledgeable, aware, and interested in the secret world of crows. In this paper, we hope that our compilation inspires young scientists to continue where others have left off in the world of avian research. This manuscript can be read as detailed background information on some of the recent research on corvids, but also has the potential to be utilized as a starting point to inspire future research.

2. Neuroscience

* Corresponding Author
marina.seidl@nwacademy.org

Advisor: Molly Sultany
msultany@nwacademy.org

Since the 1900s, our understanding of avian cognitive abilities has grown considerably—shifting from a myopic, rigid viewpoint of bird behavior as fixed and genetically inherited, to comprehending their greater cognitive flexibility. Novel discoveries in the field of crow neurobiology have shattered the black box of animal intelligence. As recently as twenty years ago, scientists believed crows processed information through simple mechanisms of input and output based on the genetic inheritance of fixed action patterns (Figure 1).



Figure 1. The Mechanical Wiring of Crow Brains; By Leah Jones Neuray

This figure depicts the input and output information processing system crows were commonly misunderstood to have. Input refers to information being collected utilizing any of the five senses. Using watercolor, this painting shows the misconception of the gears of a crow's brain working to take in information and output it. It is now understood that crows are more complex than other organisms that function this way and the simple system depicted above is a false impression.

Both fascinating and complex, the neuroanatomy of a crow's brain reveals the depth of its cognitive abilities. Crows have brain regions—such as the forebrain—comparable to primates and scientists assume these brain areas function similarly (Colombo, 2019). Just like primates, highly intelligent creatures such as crows deserve thorough scientific research and continued study. The forebrain, the region near the forehead, allows for effective information processing. Two major regions of human and crow brains include the hippocampus, which assists in learning, memory, and spatial

awareness, and the amygdala, which interprets emotions and responds to perceived threats. The parallel functionality of these brain structures reveals that crows and humans may process information in congruent ways, shedding light on the fascinating similarities in neuroanatomy (Figure 2) (Marzluff, et al., 2012, & Miyaoka, et al., 2012).

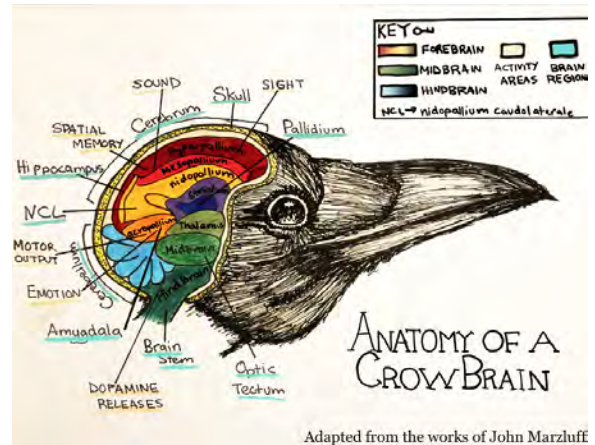


Figure 2. The Anatomy of a Crow Brain; By Marina Seidl

A colored, simplified, and concise recreation of one of John Marzluff's depictions of the anatomy of a crow's brain. This figure contains labels and a key that describe the areas of the brain in the forebrain, midbrain, and hindbrain. This marker-drawn piece demonstrates some of the key anatomical features as well as the activity areas/functions associated with the brain regions. Understanding how corvid brains are organized and how they are unique can help researchers dive deeper into their cognitive abilities.

The evolution of human and crow brains represents an example of convergent evolution due to similar shared selective pressures. Both organisms possess distinct brain regions that perform comparative functions—however, instead of a prefrontal cortex (PFC), crows have evolved a nidopallium caudolateral (NCL) (Hahn, et al., 2021). This integral structure located in the forebrain performs many of the same functions as the mammalian PFC, facilitating higher-level thinking in crows such as flexible behavior, integrative planning, and decision-making (Güntürkün, 2005). In addition, despite the absence of a neocortex, crows demonstrate many skills often attributed to this brain

region in primates: spatial reasoning, conscious thought, and self-awareness (Butler & Cotterill, 2006). The cerebellum—located in the midbrain behind the spinal cord and brain stem—manages motor control and directs voluntary movements in both humans and crows. The size of the cerebellum may directly correlate with an organism’s motor skills, including balance and physical coordination. An enlarged cerebellum may facilitate precise forelimb control, like those seen in squirrels and primates (Paradiso et al., 1997). The cerebellum controls beak movements in crows; therefore, the comparatively large size of the cerebellum may explain crows’ precise beak coordination, who utilize their beaks as a multi-functional Swiss army knife (Figure 3).

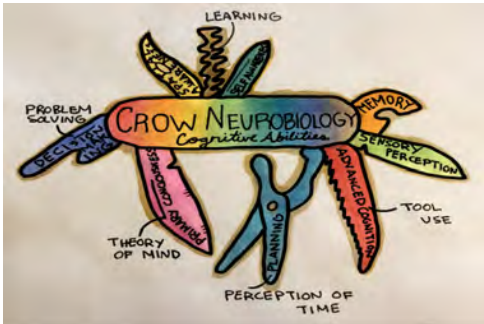


Figure 3. Swiss Army Crow; By Marina Seidl
A drawing of a swiss army knife as a metaphor for the different cognitive abilities of corvids. Crows have a multitude of abilities that aid them in their daily lives and help them survive and reproduce. Some of these depicted in the figure above include problem-solving, decision-making, primary consciousness, memory, sensory perception, planning, tool use, and more. All of these cognitive “tools” are like the different pieces of a multifunctional swiss army knife.

While primates and crows share similar abilities, including problem-solving, reasoning, and planning, the morphological organization of their brains is different. Unlike primates, who display extensive brain cortical folding, crows do not have separate brain layers. The crow brain is organized in tightly packed, dense neuronal clusters, called nuclei, with greater densities in the forebrain region (Figure 4; Olkowicz, 2016).

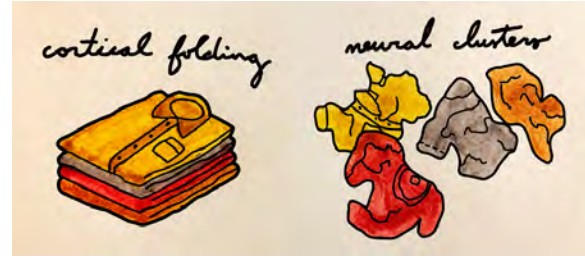


Figure 4. Laundro-clusters; ByLinnea Anderson
A drawing of laundry as a metaphor for the comparable brain structures of primates and corvids. Primates have layered brain folds, much like clean, folded, laundry. Crows, on the other hand, have tightly packed clusters, much like a pile of dirty, wrinkled, laundry. Using metaphors to interpret the anatomy of organisms can aid in the understanding of how those structures function.

Previously, scientists believed the organization and placement of specialized neurons determined a species' cognitive abilities. However, with additional research, the circuitry and the transmission of information between neurons may better explain an animal’s cognitive abilities (Shanahan, 2012). By studying the circuitry and structure of the neuron-dense pallium of crows, scientists revealed functional similarities in the way the neurons fire and communicate. Surprisingly, with smaller neural clusters and tightly packed neuroanatomy, corvids have more neurons per brain area (i.e. higher neural density) than a primate brain (Olkowicz, 2016).

Although primary consciousness, defined as external and internal awareness demonstrated by an organism (Griffin, 2003), is often considered a unique identifying trait of primates—particularly of humans—some birds and mammals may possess a similar level of self-awareness (Figure 5).

Yet in non-human animals, recognizing consciousness from a neurobiological or behavioral perspective presents a challenge. Interspecies comparisons of intelligence may be misleading, like comparing apples to oranges, with animals displaying a more advanced spatial intelligence compared to humans’ linguistic prowess (de Waal, 2016). However, research suggests that crows have a unique and sophisticated sense of self-awareness—a proxy for intelligence (Figure 6; Nieder, et al., 2020).



Figure 5. The Steady Climb; By Ingrid Lam
A watercolor piece depicting the progression of how much researchers know about corvid cognitive abilities. On the pyramid in this figure, humans are placed at the top. This demonstrates how humans are thought to be the most cognitively advanced species. The painting depicts how crows are climbing higher on the pyramid of our conceptions of organism intelligence and competence.

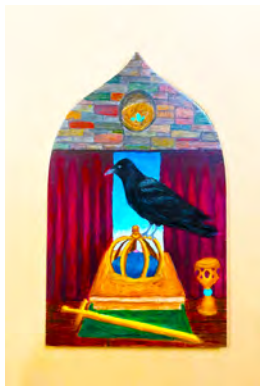


Figure 6. A Noble Bird Cage; By Ingrid Lam
An acrylic painting of a corvid perched on a crown demonstrates the resilience and courageousness of the crow. This painting is inspired by egg tempera style renaissance era paintings. Painted in a portrait style, the main foreground of the painting is a crow. This is a counterintuitive cultural depiction of crows in a positive light.

By observing neural response in the pallium region of a crow's brain during tasks involving spatial awareness and self-awareness, scientists documented that crows possess a basic level of consciousness (Nieder, et al., 2020). Recent research on crow neurobiology suggests newly defined brain structures may help explain corvids' high intelligence, including both causal reasoning and sensory consciousness (Marzluff et al., 2012).

3. Communication

Crows communicate through a wide variety of sounds using their syrinx to produce a diverse array of chortles, caws, grunts, barks, and coos (Chamberlain, et al., 1968). Known as one of the more vocal birds, crow vocalization may signify danger, food resources, or social interactions. For example, crows in the same familial group learn and mimic the calls of other individuals, gradually learning appropriate tones and phrasing for differing situations. During this period of adolescence, crows observe diverse behaviors such as caching of food, tool use, and predator-specific vocalizations (Max Planck Society, 2020). Crows utilize a traditional loud, hoarse caw sound when communicating, with variations to signal alarm, distress, or group assembly (Siriwardena, 2015). In addition, researchers have observed some regional variation in crow dialects based on geographic location (Westerfield, 2011). Regional variations in vocalizations and body language in crows resemble the four different dialects of American English spoken within the north, south, west, and midland areas of the United States (Marzluff, 2005).

Crows live in close-knit family groups, often consisting of two to eight individuals, and form strong familial bonds involving extended parental care and social learning (Figure 7; Marzluff, 2005).



Figure 7. The Crows from Springfield (Inspired by The Simpsons); By Linnea Anderson
A drawing representing the crows as the Simpsons family, showing the strong bonds of corvid families. The Simpsons family is a pop culture symbol. They are commonly known throughout America as the classic family. Crows are similar in that they stick together in their familial groups and rely on each other for communication and learning.

The rich and complex patterns of verbal and gestural communication among crows represent evidence of social learning, which is the comprehension of novel behaviors acquired through observation (Wascher et al., 2012). In addition, the horizontal transfer of information from crow parents to their offspring suggests the ability to recognize individual human faces, and convey information about the potential danger level of people passing by (Marzluff, 2010). During an extended period of adolescence, juvenile crows may assimilate skills through observation, experience, modeling, and practice. Interestingly, crow familial groups may react differently in the presence of aerial versus terrestrial predators (i.e. owl and raccoon; (Figure 8).

Predator-specific calls may encode valuable information to crows, warning the group of approaching predators. Changes in vocalization patterns suggest a perceived assessment of threats. Raccoons, for example, may hunt primarily crow chicks and eggs, as opposed to owls, who may hunt crows at any stage of their lives, therefore raccoons are assessed as a higher threat.



Figure 8. Notes of Terror; By Ingrid Lam
A watercolor piece showing the rich and complex verbal communication of crows. Especially in the presence of an aerial predator such as an owl, crows react to threats. Crows have a diverse range of sounds they make, including warning signals and alerts of potential predators. Owls and raccoons are common predators of crows and are a threat to crow families and eggs.

4. Learning

Crows can not only remember specific facial

features and behaviors of predators and threatening individuals, but they can also pass this specific knowledge on to their offspring. The mind of a crow operates like that of a chess player: capable of planning multiple steps ahead (Figure 9).

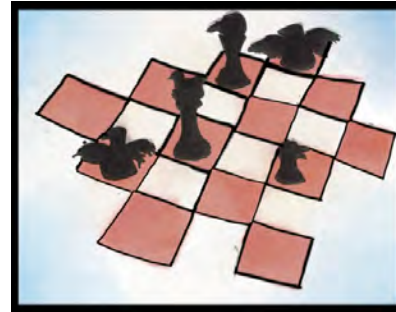


Figure 9. Checkmate; By Ingrid Lam
A figure of crows as chess pieces. Each chess piece on the board is the head and wings of a crow. Crows' minds are complex and able to plan ahead. With abilities to understand things spatially, the cognitive capacities of corvids are expansive and involve multiple layers of complex neural flexibility.

In addition, crows have a well-developed spatial and working memory; when caching, they intentionally hide food in trees or tall grass, attempting to conceal these food resources from other pilfering crows, with the goal of returning at a later time to relocate these items for consumption. Perishable food may be retrieved earlier than other food resources, demonstrating a high degree of planning and cognitive flexibility (Emery & Clayton, 2004).

With an impressive cognitive tool kit, crows demonstrate the ability to learn, recall, and utilize insight to solve complex problems in their daily lives. Their impressive spatial memory represents an adaptive advantage. Some behavioral scientists call crows flying monkeys due to their advanced, primate-like intelligence (Marzluff, 2005). In addition, crows hold funerals as a way to learn about any potential danger. Through the implementations of customary funerals and secretive caching, crows use their strong spatial and working memory as a survival skill and cultural meme (Figure 10; Swift, 2015).

Today, scientists understand that birds in the Corvus genus have sophisticated cognitive abilities.

For example, in multigenerational family groups, crows may learn how to utilize tools, the act of utilizing an object to complete an action, in pursuit of food resources. Young crows mimic and mirror specific actions while watching their adult counterparts. In New Caledonian crows, (*Corvus moneduloides*), tool use may demonstrate evidence of social learning and generate positive emotional behavior. When utilizing tools, New Caledonian crows completed a spatial problem-solving task faster, demonstrating a complex memory, and elaborate understanding of object permanence—the ability to remember an object even when out of sight (McCoy et al., 2019). Understanding motivation and goal-directed enrichment of animals—both in captivity and the wild—represents a promising avenue for future crow research.



Figure 10. The Mourning Birds; By Pritam Khalsa
A dramatized depiction of a crow's funeral in a cemetery. Crows gather around the dead to understand the dangers of the area. Often thought to be funerals, crows gather information and learn from deceased crows to avoid danger and threats. Survival skills are an observable trait that can be shown in many ways.

5. Cities: Prosperity in a Concrete Jungle

Cultural evolution occurs when one species adapts over time through social learning. A murder of crows may demonstrate the cultural transmission of knowledge in raising their young or warning others of the imminent danger of distinct predators (Max Planck Society, 2020). With the growth and

development of human societies throughout millennia, humans have influenced the types of available habitats for crows. Patterns of cultural evolution among humans, with the agricultural revolution ten thousand years ago, and, more recently, the rise of urbanized city centers, have influenced the realized niche of crow populations worldwide. Cities represent an attractive urban nesting, and foraging habitat, for crows, from the surplus of human food waste to fewer predators. At present, living in an urban matrix such as a city center provides crows with a diverse range of plentiful, nutrient-rich food resources, in addition to shelter, tools, and protection, resulting in a decrease in the area of territory utilized by urban crows (Król and Hernik, 2020). Crows' behavioral ingenuity, cognitive flexibility, and resourcefulness allow them to thrive in cities, reaching population sizes thirty times larger compared to crows living in forested wildlands (Marzluff, 2001).

Crows are generalist, opportunistic foragers, with a highly flexible and diverse diet. Optimal foraging theory suggests that crows may seek out food resources with the least search time and most efficient foraging efforts to conserve energy (Zach, 1979). Yet the year-round, stable availability of nutritionally deficient food resources within cities may negatively affect crow health (Heiss, 2009). While the diet of rural crows consists of a protein-rich array of resources such as berries, insects, small mammals, nestlings of other bird species, larvae fish, and frogs, urban crows may forage on human garbage, fast food, bird seed or pet food. Histology studies suggest higher blood cholesterol levels among urban crows compared to their rural counterparts; a reduction in the nutritional value of food resources may reduce reproduction rates, with lower nestling birth weights (Townsend, et al., 2019). However, the lower reproductive rates of urban crows may not result in lower crow population sizes, as rural areas may pose greater threats, resulting in higher mortality (Heiss, et al., 2009). Thus, differences in food accessibility, nutritional value, and threats to survival may influence the crow's lifespan. In addition to providing crows with a reliable, calorie-rich variety of food resources, cities offer a vast array of potential tools (Marzluff, 2005).

The sophisticated cognitive abilities of crows allow them to easily modify both natural and human-created objects in an urban matrix, solving everyday problems with ingenuity and ease through trial-and-error learning. For example, carrion crows (*Corvus corone*) in Japan may drop nuts in proximity to the automobile wheels or in front of cars at traffic stop lights. After the traffic resumes, the wheels drive over the nut, crushing the tough, inedible shell, and exposing the nutritious kernel (Marzluff, 2005). Over time, this behavior was transmitted among a population of crows, evidence of cultural learning among individuals.

Cities may represent a type of ecological enemy release. By protecting crows from the presence of many predators abundant in rural and forested areas—hawks, owls, eagles, raccoons, and foxes—crows may focus their efforts on reproduction, social learning, and foraging (Sorace, 2002). Predator detection, mobbing, and communication efforts, including dive-bombing potential predators, represent exhausting and energy-depleting activity. Therefore, urban crows may allocate more of their daily time budget to teaching their young, nest-building, and exploring resources in an urban matrix.



Figure 11. The Peril of Cities; By Pritam Khalsa
The dangers of urban spaces and humans are a threat to crows. This figure depicts monsters meant to represent humans terrorizing the crows in the city. Using watercolor, this painting dramatizes the threats of bustling cities to the well-being of crows. Dangers like intentional poisonings, cars, loud noises, unhealthy food waste, and more, can prevent crows from thriving in urban locations.

Although cities offer an attractive novel habitat, humans may represent a formidable threat to urban crow populations. In the Pacific Northwest of the United States, city locals in Portland, Oregon advocated for the relocation or removal of crows from the city center due to their messy droppings, loud vocalizations, and large downtown winter roosts during cold weather. In 2014 and 2018, the city utilized avitrol pellets, a lethal neurotoxin, to kill crows (Figure 11).

In 2019, after public outcry and vocal condemnation by the Audubon Society of Portland, Portland city officials banned the future use of avitrol to prevent the intentional poisoning of crow populations (Profita, 2019). This example illustrates how local governments may knowingly violate The Migratory Bird Treaty Act, which provides legal protection for crows, among other birds. Even with this protection, state legislation in the United States may legalize killing crows without a license or allow the elimination of crows if they pose a significant threat to agriculture, livestock, or tree nurseries. Therefore, urban crows may face lethal consequences if they are perceived as a nuisance or threat to human livelihoods.

6. Changing. Crows Throughout the Ages

Throughout history, crows and humans have lived in close physical proximity, developing a commensal relationship. Even before the development of major urban cities globally, crows have been closely intertwined with ancient human cultures with their iconography and artistic symbolism among diverse First Nation Indigenous tribes in the United States. For example, crows represent an integral part of storytelling, cultural regalia, and religious symbolism: crows signify wisdom to Pueblo and the Tlingit tribes, and a sacred bird and adept communicator to Sioux tribes. During ceremonial Sioux Ghost Dances, tribal members weave a crow's feathers into their clothing; once the final flood washes upon the earth, the crow feathers would safely lift tribal members, delivering them safely to the heavens. For Sioux, Tlingit, Hopi, Chippewa, and Pueblo tribes, crows represent an interesting duality: creator of the world yet also a wily, sentient trickster

(Marzluff, 2005). Recognizing the diverse artistic, symbolic, and cultural representation of crows in Indigenous cultures will foster greater awareness of this sacred bird among First Nation peoples.

Yet, humans' perception of crows has shifted throughout history, from reverence to revulsion. In Middle Ages and Renaissance Europe, crows were viewed as an omen of foreboding evil (Król & Hernik, 2020). During the American Civil War (1861-1865), crows assumed a negative connotation, representing a harbinger of death as soldiers and civilians observed crows actively scavenging on human corpses (Marzluff, 2005). At this time, literature prominently featured crows as symbols of negative prophecy, painful loss, or unimaginable tragedy. In the words of 19th-century American writer and poet, Edgar Allan Poe, the Raven represents a macabre symbol of grief: quoth the raven, nevermore. This illustrates the power of birds as a symbol of loss and tragedy and underscores their impact on human culture and society in diverse ways (Figure 12).

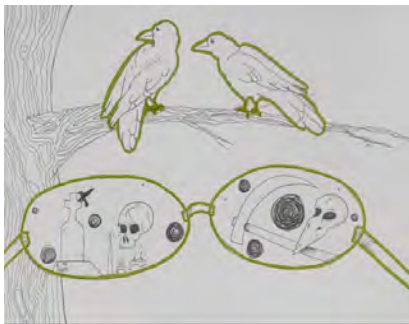


Figure 12. Through the Looking Glass of Fear; By Leah Jones Neuray

This figure shows a pair of glasses facing two crows on a branch. The glasses represent the way people negatively view crows as symbols of death. The misconceptions about crows culturally are the cause for the disregard of corvids as a valuable species. Seen in visual art, poetry, indigenous stories, and more, the culture regarding crows has seldom been positive.

In addition, 19th-century artists from American ornithologist and painter, John James Audubon (1785-1851), to Dutch post-impressionist painter, Vincent van Gogh (1853-1890), represented crows in their works. Celebrated for his detailed paintings of

crows in naturalistic habitats, Audubon depicted crows in a representational style. In European masterpieces, crows may depict hidden meanings. For example, Vincent van Gogh's iconic painting, *Wheat Field With Crows*, depicts a large expanse of wheat fields with a nearby murder of crows circling the tumultuous sky above. In this painting, one of van Gogh's last, crows may symbolize the sadness and loneliness that characterized his final days (van Gogh Museum Amsterdam, N.D).

Lastly, English language word etymology—the study of words and their cultural context and connotations—conveys human perceptions of crows throughout history. While the popular lexicon of a group of people may shift in popularity and usage over time, many historical and modern interpretations of crow biology remain in use today. For example, the English idiom, to eat crow, signifies facing humiliation or defeat by admitting one's defeat or setting aside one's pride. Interestingly, the origin of this colloquial saying may be attributed to English author, Rudyard Kipling in a short story written in 1885, as a character eats a wild crow, a metaphor for an inescapable situation. To crow over someone means to brag, gloat, or loudly delight in someone else's defeat, yet to crow at someone may reference the hoarse, piercingly loud vocalizations of crows (Figure 13; Voice of America Learning English, 2017).



Figure 13. To Crow or Not to Crow; By Marina Seidl
Watercolor painting depicting crow vocalizations in response to potential danger, in this case, a raccoon. Crows can respond to different threats using different sounds to warn other crows of predators. Although thought to be obnoxious pests with their caws and croaks, crows communicate with each other using those noises. Corvids have a complex language researchers are still trying to understand.

Despite their close relationship with humans and fascinating representation in art, language, and literature, crows are often viewed as noisy, obnoxious pests. This perception may be partly accurate, as research suggests that rural crows may shift their behavior after being shot and killed for pilfering crops or calling loudly in proximity to humans. While the Migratory Bird Act outlines guidelines for crow protection, these regulations marginally protect crows due to existing loopholes: it is legal to kill crows 124 days of the year in response to the destruction of property, wildlife, livestock, fruit or ornamental trees, nuisance, or human health hazard, with no bag limit (U.S. Fish & Wildlife, N.D.).

7. Conclusion

Overall, groundbreaking research on the behavior, biology, and cognitive abilities of crows shed new light on these intelligent vertebrates. Findings like these indicate the vast amount of information left to discover and study the behavioral functions and cognitive abilities of the crow. Although some humans still associate crows with noise, annoyance, and vexation, many learn the remarkable abilities of the bird. With continued research on crows, humans may develop a more nuanced understanding of the complex, multifaceted, and deep-rooted interactions with crows, shaping the cultural co-evolution of both humans and crows in the future. Co-evolution defines a vast majority of the crow and human relationship, representing the significance each species has in each other's lives and its continued growth throughout history. The unique layering of neurons and other anatomical brain structures allows crows to have a diverse cognitive tool kit. Corvid intelligence and the extent of their levels of consciousness have only recently begun to be explored. Researching this further can explain the depth of crowd social learning and communication. Crows share close-knit family groups and can learn from each other, a skill found in many cognitively advanced species. Other skills include tool use and the ability to adapt and thrive in urban and rural environments. These abilities can indicate to researchers the extent of corvid complexity. This paper was written to inspire future researchers to expand what's known about corvids in

the world of behavioral ecology. Corvids are avians that deserve full exploration of their cognitive abilities and complex, intriguing, culture. Their history has been constantly intertwined with humans and merits a deeper dive into the secret world of crows.

Acknowledgments

Thank you to Professors John Marzluff and Cody McCoy whose helpful comments improved this paper. We would also like to acknowledge our student copy editors at Northwest Academy: Aaron Drummond, Maddox Brewer Knight, Zadie Niedergang, and Zola Young. Special thanks to Dashiell Weinstein and Elizabeth Clarke for the digital editing of paper figures.

References

- Butler, A. B., & Cotterill, R. M. (2006). Mammalian and avian neuroanatomy and the question of consciousness in birds. *The Biological Bulletin*, 211(2), 106–127. <https://doi.org/10.2307/4134586>
- Colombo, M. (2019). Avian brains: Primate-like functions of neurons in the crow brain. *Current Biology*, 29(16): Pages R794-R796 <https://doi.org/10.1016/j.cub.2019.07.001>
- Dwight R. Chamberlain, & Cornwell, G. W. (1971). Selected Vocalizations of the Common Crow. *The Auk*, 88(3), 613–634. <http://www.jstor.org/stable/4083753>
- Emery, N. J., & Clayton, N. S. (2004). The Mentality of Crows: Convergent Evolution of Intelligence in Corvids and Apes. *Science*, 306(5703), 1903–1907. <http://www.jstor.org/stable/3839824>
- Griffin, Donald R., and Gayle B. Speck. "New Evidence of Animal Consciousness." *Animal Cognition*, vol. 7, no. 1, 1 Jan. 2004, pp. 5–18, 10.1007/s10071-003-0203-x.
- Güntürkün, O. (2005, November 2). *The Avian 'prefrontal cortex' and cognition*. Current Opinion in Neurobiology. Retrieved from <https://www.sciencedirect.com/science/article/abs/pii/S0959438805001534>

- Hahn, L. A., et al., (2021, December 3). *Working memory capacity of crows and monkeys arises from similar neuronal computations*. eLife. Retrieved from <https://elifesciences.org/articles/72783>
- Heiss, R. S., Clark, A. B., & McGowan, K. J. (2009). Growth and nutritional state of American crow nestlings vary between urban and rural habitats. *Ecological Applications*, 19(4), 829–839. <https://doi.org/10.1890/08-0140.1>
- Król, K., & Hernik, J. (2020). Crows and Ravens as indicators of socioeconomic and cultural changes in urban areas. *Sustainability*, 12(24), 10231. <https://doi.org/10.3390/su122410231>
- Marzluff, J. M., et al., D. J. (2012). Brain imaging reveals neuronal circuitry underlying the crow's perception of human faces. *Proceedings of the National Academy of Sciences of the United States of America*, 109(39), 15912–15917. Retrieved from <https://doi.org/10.1073/pnas.1206109109>
- Marzluff, J. M., & Angell, T. (2005). *In the company of Crows and Ravens*. Yale University Press.
- Marzluff, J. M., et al., (2010, January 18). *Lasting recognition of threatening people by Wild American crows*. Animal Behavior. Retrieved May 13, 2022, from <https://www.sciencedirect.com/science/article/abs/pii/S0003347209005806>
- Marzluff, J. M., & Angell, T. (2007). *In the company of Crows and Ravens*. Yale University Press.
- Marzluff, J. M., Bowman, R., & Donnelly, R. (2001). *Avian ecology and conservation in an urbanizing world*. Springer US.
- Max Planck Society (2020, June 1). *Long childhoods and extended parenting help young crows grow smarter*. Phys.org. Retrieved May 13, 2022, from <https://phys.org/news/2020-06-childhoods-parenting-young-crows-smarter.html>
- McCoy DE, et al., (N.D.). *New Caledonian crows behave optimistically after using tools*. CB. Retrieved May 13, 2022, from <https://pubmed.ncbi.nlm.nih.gov/31378612/>
- Migratory bird treaty act of 1918*. Migratory Bird Treaty Act of 1918 | U.S. Fish & Wildlife Service. (2020, April 26). Retrieved May 13, 2022, from <https://www.fws.gov/law/migratory-bird-treaty-act-1>
- 918
- Olkowicz, Seweryn, et al. “Birds Have Primate-like Numbers of Neurons in the Forebrain.” *Proceedings of the National Academy of Sciences*, vol. 113, no. 26, 13 June 2016, pp. 7255–7260, 10.1073/pnas.1517131113. Accessed 27 Apr. 2019.
- P;, N. A. W. L. R. (2020, September 25). *A neural correlate of sensory consciousness in a corvid bird*. Science (New York, N.Y.). Retrieved May 13, 2022, from <https://pubmed.ncbi.nlm.nih.gov/32973028/>
- Paradiso, S., et al., (1997). Cerebellar size and cognition: correlations with IQ, verbal memory and motor dexterity. *Neuropsychiatry, neuropsychology, and behavioral neurology*, 10(1), 1–8. Retrieved from <https://pubmed.ncbi.nlm.nih.gov/9118192/>
- Profita, C. (2020, June 3). *Portland bans toxic bird poisons on City Property*. opb. Retrieved May 13, 2022, from <https://www.opb.org/news/article/portland-oregon-toxic-bird-poisons-avicides-ban-city-property/>
- Shanahan M. (2012). The brain's connective core and its role in animal cognition. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 367(1603), 2704–2714. Retrieved From <https://doi.org/10.1098/rstb.2012.0128>
- Siriwardena, Gavin Mark. (1995): Aspects of vocal communication in the carrion crow, *Corvus corone*. University of Leicester. Thesis. <https://hdl.handle.net/2381/34233>
- Sorace, A. (2002). High density of bird and pest species in urban habitats and the role of predator abundance. *Ornis Fennica*, 79. Retrieved from https://lintulehti.birdlife.fi:8443/pdf/artikkelit/322/tiedo/79_60-71_artikkelit_322.pdf
- Stacho, M., et al., (2020). A cortex-like canonical circuit in the avian forebrain. *Science*, 369(6511). <https://doi.org/10.1126/science.abc5534>
- Swift, K. (2015). *Wild American crows use funerals to learn about danger*. Research Works Archive. Retrieved May 13, 2022, from https://digital.lib.washington.edu/researchworks/bitstream/handle/1773/33178/Swift_washington_02500_14237.pdf

Townsend, A. K., Staab, H. A., & Barker, C. M. (2020, June 3). *Urbanization and elevated cholesterol in American crows*. UC Davis. Retrieved May 27, 2022, from <https://ucdavis.pure.elsevier.com/en/publications/urbanization-and-elevated-cholesterol-in-american-crows-2>

Vincent van Gogh (1853-1890), Auvers-sur-Oise, July 1890, oil on canvas, 50.5 cm x 103 cm
Credits (obliged to state): Van Gogh Museum, Amsterdam (Vincent van Gogh Foundation)

V.O.A.L. English, (2017, August 12). *When do we 'eat crow?'*. VOA. Retrieved May 13, 2022, from

<https://learningenglish.voanews.com/a/words-and-the-ir-stories-eat-crow/3975932.html>

Wascher, C. A. F., et al., (2012, September). *You sound familiar: Carrion crows can differentiate between the calls of known and unknown heterospecifics*. *Animal cognition*. Retrieved May 13, 2022, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4417708/>

Zach, R. (1979). Shell Dropping: Decision-Making and Optimal Foraging in Northwestern Crows. *Behavior*, 68(1/2), 106–117. <http://www.jstor.org/stable/4533944>

Brainwave Music: A Scientific Review of Psychological Effects and Future Innovations

Sihe Wang^{1*}

¹Beijing No.4 High School International Campus, Beijing China

Received July 31, 2022; Revised February 7, 2023; Accepted, March 9, 2023

Abstract

This literature review introduces brainwave music (BWM), a revolutionary type of music generated using digitally filtered brainwaves as displayed in electroencephalography (EEG) impulses. The review summarizes the previous studies on the applications of BWM in different fields and presents a vision for future innovative uses. The review also explains how EEG data is used to map mental state to music with similar arousal levels, and how BWM can be used in clinical or psychological settings. Additionally, the review discusses how brainwave music can have positive effects on various mental states such as anxiety, sleep, and more. The review also highlights the importance of portability and operability in future brainwave music applications, and the potential for incorporating visual elements such as video and flashes to enhance the experience. Overall, brainwave music combines neurology, science, and music therapy to improve focus, energy, stress, anxiety, pain, and sleep.

Keywords: Brainwave Music, Brainwave, Music Therapy, EEG Study, Disease

1. Introduction

For the past three years, the COVID-19 epidemic has continued to affect all aspects of the world. For the student population, the change in the learning environment from campus to home is certainly a challenge. It is an urgent need for students to be able to maintain the same state of learning at school

and to avoid the limitations of online teaching (e.g., teachers are forced to reduce their supervision) that can affect their learning outcomes. It has been observed that during the COVID-19 epidemic's home study, prolonged internet usage for leisure and inadequate learning surroundings that caused constant interruptions were factors on students' learning. In response to this

* Corresponding Author
julie.wangsihe@gmail.com

Advisor: Zijin Li
lzijin@ccom.edu.cn

need, brainwave music lists and videos, which are claimed to be one of the approaches that are able to "promote concentration and improve learning efficiency", are popular among users. Compared to medication, it is non-invasive and does not have side effects. In contrast to exercise or nutrition, it often provides immediate effects. Brainwave music (BWM) is also more portable and versatile than traditional methods, such as study groups or tutors, as it can be used anytime, anywhere, and for multiple purposes. Overall, brainwave music offers a unique and effective solution for those seeking to enhance their cognitive abilities. This literature review will introduce brainwave music and summarize the applications of brainwave music in different fields based on previous studies. Finally, the paper will present a vision of future innovative applications of BWM.

2. Definition

BWM is a revolutionary type of music generated using digitally filtered brainwaves, as displayed in electroencephalography (EEG) impulses, in accordance with certain translation criteria. (Huang et al., 2016) The brainwave is described as an arhythmic electric potential between nerve cells, also known as neurons, and is expertly caught by EEG apparatus (Rodgers, 2011). EEG is the recording of electrical activity generated by the firing of neurons in the brain along the scalp. In clinical situations, EEG refers to the

recording of the spontaneous electrical activity of the brain during a specified time period (Cohen, 1996).

3. History

In the second part of the 19th century, the physiologist Etienne-Jules Marey pioneered the graphic way of capturing diverse physiological data, ushering in the era of imaging methods. Hans Berger, a German psychiatrist, presented the first study on human EEG in 1929 by capturing scalp-layer electrical waves with a non-invasive probe.

Edgar Adrian, a renowned neurophysiologist and Nobel winner, duplicated the experimental results of Berger's rhythm in 1934 by translating EEG impulses to sound signals and playing them via a loudspeaker. What is its sound? Similar to audio signals, brain waves are frequency-varying sequences of time. Humans make sounds in the frequency range of 20 to 3000 Hz, primarily in the 100-300 Hz region, while the vast majority of musical instruments create hundreds of hertz notes. In contrast, scalp-collected EEG signals are mostly concentrated in the 1-100 Hz frequency range, and their power spectrum decays fast with increasing frequency. If the EEG signal is immediately played as an audio file, it sounds like an inconsequential low-frequency rumbling. In the beginning, due to the limits of signal processing technology, the EEG signal frequency could only be boosted by multiplying by a particular amount to make

it audible to the human ear.

Beginning in the 1950s, developments in signal processing led to considerable improvements in the analysis of EEG data, establishing the groundwork for the production of brainwave music. The world's first brainwave music was made in 1965.

Music for the single performer, the world's first brainwave music, was launched in 1965 and performed throughout Europe and the United States. Edmond Dewan, a prominent physicist, and Alvin Lucier, an experimental musician, collaborated to build a system that employed alpha waves in the brain to control the sound of percussion instruments, which were produced and performed by Alvin Lucier's brain waves. The Musical Instrument Digital Interface (MIDI) was established in the 1980s in order to tackle the problem of communication between electro-acoustic instruments (Mangieri, 2022). Towards the close of the twentieth century, advancements in EEG and sound technology led to the construction of the Brain-Computer Music Interface (BCMI), which, for instance, controls the playing of complex instruments using EEG. By mapping the signal properties of EEG to pitch, timbre, and tone in music and recompiling them into MIDI format, brainwave music became more accessible and agreeable to the ear.

4. Current Theory

Scale-free BWM formed from EEG data according to the power law of both EEG and music has the features of both music and

EEG and may contain physiological information that music alone may not (Wu et al., 2010, 2014; Lu et al., 2012).

As it has been reported that music with specific pitch or rhythm characteristics effectively promotes brain function (Adalarasu et al., 2011), perhaps in a manner similar to sound resonance, BWM created by relaxed-state EEG signals may also induce such activities in the brain and prompt specific brain function, resulting in an improved mental or emotional state (Fedotchev and Radchenko, 2013).

Table 1: Frequency Range - State of Mind

Beta 12Hz-38Hz	Wide awake state
Alpha 8Hz-12Hz	Awake but relaxed.
Theta 3Hz-8Hz	Light sleep or extreme relaxation.
Delta 0.2Hz-3Hz	Deep, dreamless sleep.

Table 1 describes different brain wave frequencies measured by EEG and the states of consciousness associated with them. Beta waves, which have a frequency range of 12Hz-38Hz, are associated with the wide awake state. Alpha waves, which have a frequency range of 8Hz-12Hz, are associated with being awake but relaxed. Theta waves, which have a frequency range of 3Hz-8Hz, are associated with light sleep or extreme relaxation. Delta waves, which have a frequency range of 0.2Hz-3Hz, are associated with deep, dreamless sleep.

The resonance of the head caused by the sound waves generated by music generates

alpha brain waves that can loosen up the mood, have self-healing ability, have a high immune status, and slowly release morphine in the brain. These waves, which are referred to as healthy waves, are capable of promoting self-healing and have a high immune status. Pang Jun found in "Modern medical clinical non-pharmacological therapy of insomnia profile" based on studies and clinical performance that activation of lower frequencies, light, and rhythmic sound induces a calm state in human brain waves.

In "Overview of Clinical Non-pharmacological Treatment of Insomnia in Modern Medicine," it was concluded from the experimental and clinical performance that stimulation of low frequency, light, and rhythmic sound induces patients to reach a state of relaxation and eventually sleep (Pang, 2005). In "On the Psychotherapy Function of Music," it is said that music therapy has evolved from a tool for entertainment to a tool for enhancing mental state and quality of life with psycho-therapeutic functions (Ge, 2006).

5. General use

Brainwave Entrainment (BWE) is a technique that uses auditory tones, flashing lights, or a combination of both, in conjunction with isochronic, monaural, or binaural beats, to stimulate the brain at the correct frequency. Isochronic tones are evenly spaced tones that alternate between being on and off. Monaural and binaural

beats are delivered as two tones with very similar frequencies, and the brain detects a beat corresponding to the pitch difference between the two tones. The pitches are delivered simultaneously with monaural beats, but each ear is supplied with binaural beats individually. Therefore, when listening to brainwave music containing binaural beats, it is recommended to wear headphones or similar music playing devices to achieve the best relaxation effect.

The process of creating music using BWE is illustrated in Figure 1, with the blue arrow representing the conceptual framework and the yellow arrow representing the actual implementation in this study. Using arousal as a bridge, EEG data is collected as a reflection of mental state and mapped to the parameters of music with similar arousal levels, according to a two-dimensional model of mood. The creation of music consists of five steps, as shown in Figure 2: 1) extracting EEG signal characteristics, 2) defining music segmentation (factors: main note, tone, and rhythm cadence) based on relevant EEG features, 3) building music bars (parameters: chord and note location) based on EEG features and song segment parameters, 4) setting properties of notes (timbre, scale, duration, and loudness) based on bar parameters, and 5) constructing the music melody using a program. The choice of feature extraction technique can affect the meaning of the resulting music, and any rule shared by brain function and music can serve as a suitable bridge between brainwaves and music. (Wu et al., 2010)

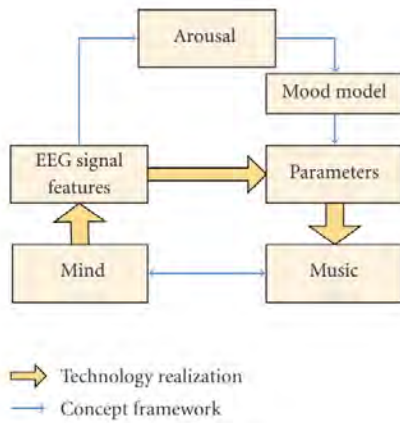


Figure 1: Overview of the brainwave music generation (Wu et al., 2010)

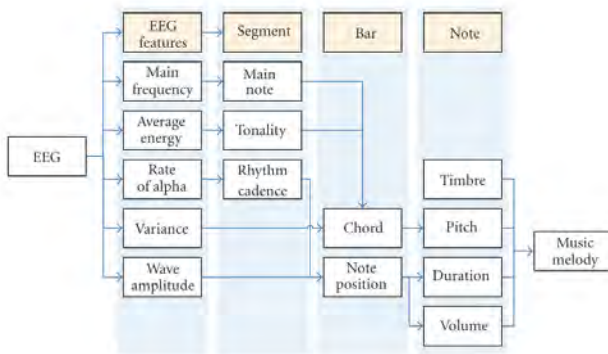


Figure 2: Mapping rules from EEG to music (Wu et al., 2010)

The technical level now attained in computer-assisted imaging elaboration as well as in the employment of novel methods of analysis has sparked significant fresh interest in both the theoretical analysis of the therapy and its clinical applications.

6. Methods

In order to conduct a thorough review of the existing literature on brainwave music, a comprehensive search strategy was employed, and total 50 works were eventually reviewed. A total of ten databases

were utilized, including WorldCat, PubMed, Paddyfield, CampusWell, SageJournal, ESource, CNKI, Wanfang database, ScienceDirect, and Frontiers. The search was limited to articles published in English, with a few exceptions for articles written in Chinese and German. The search terms used included *brainwave music*, *music therapy*, *stress management*, *attention performance*, *neural entertainment*, *EEG analysis*, *music stimuli*, *emotion perception*, *musical emotion*, *neural plasticity*, *binaural beats*, *monaural beats*, *isochronic beats*, *isochronic tones*, *brainwave entertainment*, *brain stimulation*, *photic stimulation*, *audio-visual entertainment*, *AVE*, *brain AND entertainment*, *afferent sensory stimulation*, *sleep*, and *sleepiness*.

The selection criteria for the articles included in this review were as follows: (1) the article must address brainwave music or a related area of expertise; (2) the article must include an experimental study with a comparison group or a pre-test/post-test design; (3) the study must assess clinical or psychological outcomes using valid and adequate test techniques; (4) the study must provide enhanced statistical data, including descriptive statistics and analysis, as well as figures and tables to illustrate the outcomes.

In addition to the articles retrieved through the literature search, a song list of brainwave music provided on the network platform was also chosen to test the application of brainwave music at this stage. A total of 20 brainwave music playlists were selected from three audio-video platforms: QQ music, Netease cloud music, and

Youtube. After excluding the overlapped singles, ten complete playlists were compiled, with a total of 182 audio tracks.

The number of samples collected in the studies included in this review varied, and the filtration method used to narrow down the key resources adopted for the research was not specified in all of the studies. However, the selection criteria outlined above were used to ensure that only high-quality, relevant studies were included in this review.

7. Results

7.1 Cognitive

Multiple types of research have demonstrated that musical compositions and deep brain stimulation have a beneficial facilitative influence on people's cognitive behavior. The primary areas affected were attention, productivity, cognitive ability, and academic achievement. Five of the research were done on children and adults with attention deficit hyperactivity disorder (ADHD). At the same time, the remaining investigations were conducted on college students or persons employed in a particular field.

According to studies using modified visual oddball tasks (Kilmesch et al., 1998), desynchronization in the lower alpha band reflects attention, whereas desynchronization in the upper alpha band represents semantic memory performance. Theta band synchronization also reflects episodic memory and the encoding of new

information (Buzsaki et al., 1994). Two studies examined the effect of photic stimulation with or without electrodermal stimulation (EDR) at various frequencies on general intelligence or grade point average (GPA). The research recruited youngsters with ADHD or college students with "academic problems" and found that stimulation of 14 Hz alternating with 22 Hz or 12 to 14 Hz over numerous sessions increased GPA or the Scholastic Achievement Wechsler Individual Achievement Test substantially.

In addition, research indicates that rhythmic classical music might reduce the tension associated with number crunching and lead to more reasonable customer decisions. (Feng et al., 2014). Math-averse consumers shun alternatives requiring pricing calculations. Intriguingly, these customers experience less arithmetic anxiety when classical music with a moderate pace is playing in the background. Nonetheless, in the absence of background music or when the speed of such music is rapid, people with high math anxiety avoided decisions involving pricing computations more.

The influence of background music on listeners' attention test scores increases with the intensity of listeners' perception of background music relative to the absence of background music, according to studies. (Huang & Shih, 2011) Therefore, lyrics and melodies that tend to attract the listener's attention should not be included in music used for work or study. Brainwave music, as pure music with a tranquil melody and strong rhythm, successfully circumvents the

ban on working background sound and may play a more significant role in stimulating thought.

Evaluation of the influence of an external element, particularly music playing, on the execution of a task requiring significant concentration. (Mori, 2014) The result suggests that listening to the favorite music condition improves overall performance since it decreases the number of instances in which the subject is unable to locate the exceptional indication and takes an abnormally long time to complete the trial.

7.2 Stress and Anxiety

Brainwave music has substantial impacts on compulsiveness, interpersonal relationships, sadness, anxiety, anger, paranoia, and psychoticism. It was determined that there was no interaction between the gender factor and the effect of brainwave music; it was also determined that the effect of brainwave music on the emotional state of college students was significantly different in the music cultivation dimension; based on these results, it was determined that there was no interaction between the gender factor and the effect of brainwave music.

In the relaxation curve experiment, the brainwave music group demonstrated a progressive increase in the overlap. In contrast, in the control group (refers to the Youth Digest and French News group), there was no discernible trend of change. In conclusion, the regularity and consistency of the integrated relaxation curve in the

brainwave music group suggests that brainwave music may induce a calm state in each experiment's individuals.

In addition, researchers have discovered that healthy persons may utilize music to reduce stress through music creation, drumming, and passive listening. It can also provide essential assistance for physical activity (Parambi & Prabhakar, n.d.).

The research on the extended length of brain functional cortex (FC) in response to music, network density, and the number of brain connections in the Music Listening Group revealed substantial changes not only in the general connectivity but also in specific frequency bands, such as the delta, alpha, and beta bands. This alteration implies an improvement in the subject's level of relaxation, attentiveness, mental exertion, and awareness. (Mahmood et al., 2022) Long-term research on the effect of music listening on the human brain reveals that the FC in the frontal and parietal lobes within the alpha and theta frequency bands significantly increases, while the FC in the frontal, temporal, and occipital lobes within the beta frequency band significantly decreases. This rise in alpha and drop in beta FC shows that the individual has been relaxed (Kim et al., 2014).

Brainwave entertainment employing binaural beats significantly reduced state anxiety and trait anxiety in 30 engineering research firm employees who listened to non-vocal music for four weeks, three times per week. Every session lasted around 20 minutes. This indicates that brain wave coordination through alpha and beta sounds

in relation to stress condition and stage is effective and can be used as a non-invasive method for controlling and reducing stress among patients in clinics and treatment centers, by psychologists, counselors and psychiatrists. (Alipoor et al., 2014).

7.3 Pain

In 2016, researchers studied the impact of brainwave music on alleviating orthodontic discomfort. 36 individuals matching for age, gender, and anxiety/pain levels were randomly allocated to either the BWM (N=12), CBT (N=12), or control (N=12) groups (Huang et al., 2016). The BWM group's baseline resting EEG data were transformed into BWM. During the first week following orthodontic appliance insertion, EEG signals and pain perception (as measured by the Visual Analog Scale, VAS) were collected for each participant. Multiple methods were used to evaluate EEG data.

Ultimately, the BWM and cognitive behavioral therapy (CBT) groups had significantly reduced pain perception compared to the control group, demonstrating that BWM is beneficial in managing orthodontic pain, presumably by restoring functional connections and brain regularity impacted by pain (Huang et al., 2016).

In a different research, participants with bruxism and myofascial pain dysfunction syndrome were given isochronic tones of constant frequency and duration that were modified and chosen by the patient, as well

as electromyographic (EMG) feedback (Manns et al., 1981). The subjects noticed much-reduced discomfort and muscle spasms in the temporal mandibular joint after three weeks.

7.4 Sleep

A 2020 study demonstrated that slow-wave sleep brain-wave music may have a good impact on sleep quality, whereas rapid eye movement (REM) brain-wave music or white noise may not (Gao et al., 2020). Reducing the power spectral density of the delta band of EEG and increasing the functional connection between the left frontal lobe and the left parietal lobe may also contribute to improved sleep quality. As a form of musical stimulation, researchers created two forms of scale-free brain-wave music, one from the REM stage and the other from the SWS stage. Deep sleep can predict sleep satisfaction and is a representative indicator of sleep quality (Riedel and Lichstein, 1998), so for EEG analysis, we primarily analyzed the power spectrum of EEG during deep sleep and investigated the neural mechanisms of this two brain-wave music on sleep promotion from an EEG standpoint.

Another 2014 study intends to determine if the sleep quality of elite athletes may be enhanced by auditory brainwave entertainment and if this leads to improvements in post-sleep psychophysical states (Abeln et al., 2014). In preliminary research, 15 teenage top soccer players slept with binaural beats between 2 and 8 Hz for

eight weeks. Only the intervention group saw a substantial improvement in subjective assessments of sleep and waking quality, tiredness, and motivational mood, although this did not affect their reported physical condition. In conclusion, eight weeks of auditory stimulation with binaural beats increased athletes' perceived sleep quality and post-sleep condition. In contrast, the effect on the physical level is assumed to be delayed.

A 2019 study revealed that binaural beats paired with Autonomous Sensory Meridian Response (ASMR) and calibrated to the optimal frequency (6 Hz binaural beats and ASMR triggers) have the most significant sleep-inducing effects and send the user to a state of psychological relaxation the quickest. (Lee et al., 2019)

8. Limitation

The current body of research on the impact of brainwave music on cognitive function is constrained by a number of methodological considerations. One of the most significant limitations is the lack of standardization in the techniques employed to generate and administer brainwave music. Studies have employed a wide range of frequencies, amplitudes, and durations of stimulation, which makes it challenging to make cross-study comparisons. Furthermore, the sample sizes of many studies are relatively small, which can restrict the generalizability of their findings.

Another limitation is the absence of a widely accepted methodology for assessing

the effects of brainwave music on the brain. Studies have employed a range of neuroimaging techniques, such as EEG, fMRI, and PET, which can make it difficult to compare results across studies.

Finally, a paucity of long-term studies exists on the effects of brainwave music on cognitive function. Most studies have focused on short-term effects, and it remains unclear how these effects may change over time.

Despite these limitations, the available literature suggests that brainwave music holds promise as a tool for modulating brain activity. However, further research is needed to more fully understand the effects of brainwave music on cognitive function and to develop effective protocols for its use.

9. Future Use

As a burgeoning field of study, brainwave music has future research depth and application potential. Currently, brainwave music is predominantly used in audio tracks produced on music platforms and videos released on video platforms (e.g., youtube). If brainwave music is expanded to offline or immersive experiences, it will contribute considerably to the fusion of art and science.

Thus far, research has investigated the visualization of alpha brainwave music (Yu, 2020). Several methods for enhancing the immersive presentation of alpha-brainwave music were suggested by the research.

Immersive experience refers to the total immersion of the person, giving them the impression that they are in the virtual world.

Virtual reality expression is characterized by the immersive experience; immersion is the process by which the brain is impacted. It achieves the audience's conscious acceptance of more emotional input to the current events by shortening the spatial distance between the viewer and the displayed body, which is a kind of total attention in a state of forgetfulness and mental engagement, from experiencing the course of development, change, and transition therein. Virtual reality subverts the adage "seeing is believing" and breaks the incoherence of the senses created by the conventional medium of information transmission. Virtual reality creates illusions of reality through technological techniques, activating the viewer's entire body organ perception.

In recent years, as new media art has evolved, an increasing number of artworks have utilized immersive experience exhibitions to bring the audience closer. The use of surround projection and L-shaped folding projection enhances the sense of immersion in the exhibition hall, and the atmosphere created by isolating the audience from the outside world causes them to forget their original identities and enter the scene space to participate in the experience. The genuine effect of the immersive displays "TeamLab Future Amusement Park," "Feeling is Real," "I Have a Dream," "Rain Room," "Datamatics," and "Leonardo da Vinci" is warmly appreciated by the audience. The visualization of alpha brainwave music and the immersive display are constructed in public spaces or art

treatment facilities for persons under stress.

The objective of VR technology is to replicate a virtual environment in three dimensions on a computer. The user then enters the virtual realm using specialized sensing equipment to produce a novel perceptual experience. Users of VR devices may experience 3D space through visualization, 3D surround sound, and real-time interaction with the virtual world. Using this technology, an Alpha brainwave music visualization design can produce a personal immersion experience; portability and operability are more important than in the case of a physical immersion experience.

In addition to the use of brainwave music in public art facilities, which can play an anxiety-relieving and calming role, it is also possible to innovate on existing video and audio formats to create an "immersive" sense of the scene. After collecting "concentration aid" videos from different video platforms, it was found that not all videos with the term "brainwave music" in the title were using regular and effective brainwave music. In addition, the video images are relatively single, mostly illustrated animation or landscape pictures. In future use, the video images can be combined with more visual stimulation methods used in deep brain stimulation experiments, such as regular flashes, and with playback devices such as headphones, to achieve simple brain wave activation. The changing learning environment of home learning has led to a greater reliance on such artificial tools to aid concentration and sleep, so it is also important to take into

account the freshness of the video format to avoid becoming immune to such stimuli through repeated use.

Using the principle of neurofeedback, a personal and unique induction learning environment is customized to help users regulate their mental state and promote learning effectiveness. For example, do cross-border commercial cooperation with time management and efficiency-enhancing APP applications to produce similar focus-inducing devices and promote them in the market so that concentration applications provide more than just white noise.

However, BWM faces some limitations and challenges in its implementation and widespread use. One of the main challenges is the technology required for EEG data recording and analysis, which can be expensive and difficult to use, making widespread implementation a challenge. Another challenge is the lack of standardization in the mapping of EEG data to music, making it difficult to compare results across different studies and applications. Additionally, while there have been some studies suggesting the potential benefits of brainwave music, more rigorous and extensive research is needed to establish its efficacy and determine the best methods for using it. Personalization is also a challenge, as EEG signals can vary greatly between individuals, making it difficult to create a one-size-fits-all solution for brainwave music. Furthermore, brainwave music is a relatively new and unconventional approach, and it may take time for it to be widely accepted and

adopted. Overall, it is important to approach the development and implementation of brainwave music with caution and to carefully consider these limitations and challenges.

10. Conclusion

Brainwave music, as a scientific product that covers a variety of cognitive science, music therapy, neuroscience, and other fields of knowledge, will be a tool that cannot be ignored in the future to regulate both physical and mental states. It has been proven to prolong and improve concentration, enhance mental energy, reduce stress and anxiety, reduce pain, and induce sleep. Nearly fifty selected papers have conducted scientific and systematic experimental studies on brainwave music or various research areas related to it (e.g. music therapy) to confirm the relevance of these indicators. In the context of the increasing pursuit of virtual context in modern times, future applications of brainwave music may focus on creating a "sense of context" to serve the user in a way that constructs a comfortable space, either real or virtual. The future application of brainwave music will be in a wide range of situations, including classrooms, libraries, wards, offices, fitness centers, airplanes, etc.

References

Abeln, V., Kleinert, J., Strüder, H. K., & Schneider, S. (2014). Brainwave entertainment for better sleep and post-sleep state of young elite soccer players – A pilot study. *European Journal of Sport*

Science, 14(5), 393–402.

Adrian ED, & Matthews BH. (1934). The Berger rhythm: potential changes from the occipital lobes in man. *Brain*, 57(4), 355-385.

Alipoor, A., Oraki, M., & Sabet, M. Y. (2014). Efficiency of brainwave entertainment by binaural beats in reducing anxiety. *Journal of Kermanshah University of Medical Sciences*, 18(1), Article 1.

Casciaro, F., Laterza, V., Conte, S., Perialice, M., Federici, A., Todarello, O., Orsucci, F., & Conte, E. (2013). Alpha-rhythm stimulation using brain entertainment enhances heart rate variability in subjects with reduced HRV. *World Journal of Neuroscience*, 03(04), 213–220.

Curtis, D. (2022). *BINAURAL BEATS, BRAIN WAVE entertainment AND THE HEMI-SYNC PROCESS*.

D. Cohen, *The Secret Language of the Mind*. London: Duncan Baird Publishers, 1996

Doherty, C. (n.d.). *A comparison of alpha brainwave entertainment, with and without musical accompaniment*. 44.

Feng, S., Suri, R., & Bell, M. (2014). Does Classical Music Relieve Math Anxiety? Role of Tempo on Price Computation Avoidance: MATH ANXIETY AND MUSIC TEMPO. *Psychology & Marketing*, 31(7), 489–499.

Gao, D., Long, S., Yang, H., Cheng, Y., Guo, S., Yu, Y., Liu, T., Dong, L., Lu, J., & Yao, D. (2020). SWS Brain-Wave Music May Improve the Quality of Sleep: An EEG

Study. *Frontiers in Neuroscience*, 14, 67.

Geretsegger, M., Elefant, C., Mössler, K. A., & Gold, C. (2014). Music therapy for people with autism spectrum disorder. *Cochrane Database of Systematic Reviews*, 2016(3).

Huang, R., et al., (2016). The effects of customised brainwave music on orofacial pain induced by orthodontic tooth movement. *Oral Diseases*, 22(8), 766–774.

Huang, R.-H., & Shih, Y.-N. (2011). Effects of background music on concentration of workers. *Work*, 38(4), 383–387.

Huang, T. L., & Charyton, C. (2008). A comprehensive review of the psychological effects of brainwave entertainment. *Alternative Therapies in Health and Medicine*, 14(5), 38-50

Jackson, N. A. (2003). A Survey of Music Therapy Methods and Their Role in the Treatment of Early Elementary School Children with ADHD. *Journal of Music Therapy*, 40(4), 302–323.

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. *Brain Research Reviews*, 29(2–3), 169–195.

Knapp, R. B., & Lusted, H. S. (1990). A Bioelectric Controller for Computer Music Applications. *Computer Music Journal*, 14(1), 42.

Lee, M., Song, C.-B., Shin, G.-H., & Lee, S.-W. (2019). Possible Effect of Binaural Beat Combined With Autonomous Sensory Meridian Response for Inducing Sleep. *Frontiers in Human Neuroscience*, 13, 425.

- Mackenzie, M. (2021, November 10). How binaural beats, pink noise, and ASMR can help you sleep better. *CampusWell*.
- Madjar, N. et al., (2020). Contrasting effects of music on reading comprehension in preadolescents with and without ADHD. *Psychiatry Research*, 291, 113207.
- Mahmood, D., Nisar, H., Yap, V. V., & Tsai, C.-Y. (2022). The Effect of Music Listening on EEG Functional Connectivity of Brain: A Short-Duration and Long-Duration Study. *Mathematics*, 10(3), 349.
- Mori, F. (2014). The Effect of Music on the Level of Mental Concentration and its Temporal Change: (2014). *Proceedings of the 6th International Conference on Computer Supported Education*, 34–42. <https://doi.org/10.5220/0004791100340042>
- Okawa, H., Suefusa, K., & Tanaka, T. (2017). Neural entertainment to Auditory Imagery of Rhythms. *Frontiers in Human Neuroscience*, 11, 493.
- O'Reilly, R. C., & Munakata, Y. (2000). *Computational Explorations in Cognitive Neuroscience: Understanding the Mind by Simulating the Brain*. MIT Press.
- Patrick, G. J. (1996). Improved Neuronal Regulation in ADHD: An Application of 15 Sessions of Photic-Driven EEG Neurotherapy. *Journal of Neurotherapy*, 1(4), 27–36.
- Schack, B., Klimesch, W., & Sauseng, P. (2005). Phase synchronization between theta and upper alpha oscillations in a working memory task. *International Journal of Psychophysiology*, 57(2), 105–114.
- Smith, C. A., & Morris, L. W. (1977). Differential Effects of Stimulative and Sedative Music on Anxiety, Concentration, and Performance. *Psychological Reports*, 41(3_suppl), 1047–1053.
- The Nobel Prize in Physiology or Medicine 1932*. (n.d.). NobelPrize.Org. Retrieved June 18, 2022, from www.nobelprize.org/prizes/medicine/1932/adrian/lecture/
- Wagner, M. J. (1975). Effect of Music and Biofeedback on Alpha Brainwave Rhythms and Attentiveness. *Journal of Research in Music Education*, 23(1), 3–13.
- Wu, D., Li, C., Yin, Y., Zhou, C., & Yao, D. (2010). Music Composition from the Brain Signal: Representing the Mental State by Music. *Computational Intelligence and Neuroscience*, 2010, 1–6.
- Yu S. Miao. Visualization of alpha brainwave music in stress-relieving music. (Doctoral dissertation, Dalian University of Technology).
- Zhang, F., Liu, K., An, P., You, C., Teng, L., & Liu, Q. (2017). Music therapy for attention deficit hyperactivity disorder (ADHD) in children and adolescents. *Cochrane Database of Systematic Reviews*.
- Zulkurnaini, N. A., Kadir, R. S. S. A., Murat, Z. H., & Isa, R. M. (2012). The Comparison between Listening to Al-Quran and Listening to Classical Music on the Brainwave Signal for the Alpha Band. *2012 Third International Conference on Intelligent Systems Modelling and Simulation*, 181–186.

Liquid-Liquid Extraction Method for Low-cost, Low-energy Desalination of Seawater and Brackish Water

Deven Chakrabarti¹*

¹Moorestown High School, Moorestown, NJ USA

Received October 28, 2022; Revised January 19, 2023; Accepted, January 25, 2023

Abstract

More than a third of the world's population lives in countries that use more than 40% of their available supply of freshwater each year, an unsustainable level of demand. Most such countries have access to a sustainable water resource: seawater. However, desalination is currently too expensive for many of the communities that need it most. This paper presents a new desalination method: solvent-based liquid-liquid extraction that exploits the selective solubility of freshwater over saltwater in a particular class of polar organic solvents at certain temperatures below 100°C. For once such solvent, *N*-cyclohexyl pyrrolidone, the calculated desalination efficiency at 3 hours was 99.6%. Optional purification with a standard material yielded acceptably pure water. The method holds promise for development as a simple, low-energy, low-cost, low-waste process that can make desalination feasible in a broader range of coastal and marine applications.

Keywords: Desalination, Liquid-liquid extraction, Seawater, Brackish water, Polar organic solvent, Sustainable water supply

1. Introduction

Of the vast amount of water on Earth, only a tiny fraction—about 1.1%—is found in surface sources and aquifers, according to observations made in 2002–2008 (Trenberth, et al., 2011). Those sources are inequitably distributed, and many regions of the world struggle with inadequate supplies of freshwater for household, agricultural, and industrial use. According to the World Research Institute, 17 countries experience “extremely high” water stress. That is, they typically withdraw more than 80% of the available supply annually (Hofste, et al., 2019), leaving little margin for extreme events. These countries represent one-quarter of the world's population. When the criterion is merely “high” stress—that is, 40% or more of available supply is

withdrawn annually—the number of at-risk countries jumps to 44, home to one-third of the global population (Hofste, et al., 2019). The issue is not just drinking water: according to the United Nations, 69% of global water withdrawal is used for agriculture; regionally that rate can be far higher, e.g., in Africa and Asia (Food and Agricultural Organization, 2020; Gleick and Cooley, 2021). Much of this water is being drawn from nonrenewable groundwater reserves at an unsustainable rate (Bierkens and Wada, 2019).

Most at-risk countries have easy geographical access to a potential solution: desalination of saltwater. Industrial-scale desalination was essentially nonexistent before the mid-1970s but expanded rapidly after 2000 (Gleick and Cooley, 2021, Fig. 10), as technologies improved and costs declined. The

* Corresponding Author
devenchakrabarti@comcast.net

Advisor: Sean Watson
swatson@mtps.us

decline in cost has been just enough for affluent nations to access desalination. Today there are approximately 15,900 desalination facilities in operation, producing about 95 million cubic meters of usable water per day; about 48% of that output is produced in North Africa and the Middle East (Jones, et al., 2019). However, with current technology, desalination plants are both expensive to build and expensive to operate. As a result, this plentiful resource is out of reach for many of the people who need it.

The search for alternatives includes extraction approaches. One emerging method is based on nonionic surfactant extraction (Chakrabarti, 1997). This method shows good desalination efficiency but has limitations in the context of producing drinking water at scale: the water yield is low given the amount of nonionic surfactant used, the kinetics of separating the freshwater are too slow, and removing trace surfactant from the freshwater is difficult. The objective of this work was to overcome these limitations by using a solvent extraction method based on a principled choice of solvent (section 1.2).

1.1 Desalination Methods

Desalination refers to removal of salt from seawater or brackish water, rendering it usable for human activities. Typical seawater has a salinity of about 3%–3.5% salt (30–35 g/kg or parts per thousand [ppt]). The salinity of brackish water (typically found where freshwater meets seawater and in some groundwater reserves) is about 0.05%–2.9% (0.5–29 ppt).

The two most common technologies used for industrial desalination are reverse osmosis and distillation, at 69% and 25% of global capacity, respectively (Jones, et al., 2019, Fig. 5). In the reverse osmosis method, water is pumped through a semipermeable membrane that blocks dissolved salts but allows water to pass through. Distillation involves bringing seawater to its boiling point, heating further to convert water into vapor, then condensing the vapor back into water.

Both processes are costly. In reverse osmosis, the water is pumped through the membrane at high pressure, typically 800–1000 pounds per square inch,

which requires significant energy; operating the high-pressure pumps accounts for about 25% - 40% of the cost of water produced by reverse osmosis (LiVecchi, et al., 2019). Because of high pressures and sophisticated membrane technologies, reverse osmosis plants are also expensive to build. The ballpark figure has historically been \$1,100 per $1 \text{ m}^3/\text{d}$ capacity, though that is dropping as technology improves (Gasson, 2020). Thus, construction costs for a plant producing $600,000 \text{ m}^3/\text{d}$ might be close to \$700 million dollars. For perspective, in Libya, a coastal country under “extremely high” water stress, such a plant could replace about 10% of the 6.5 million cubic meters drawn daily from desert aquifers by the Great Man-Made River Project (Chibani, 2022). But construction would cost roughly 1% of the country’s gross domestic product, which is about \$48.7 million (PopulationU, n.d.).

In thermal distillation, significant energy is required first to heat the water to boiling and then to drive the phase change from liquid to vapor. As result, a thermal distillation plant uses on the order of 10 times more energy than a typical seawater reverse osmosis plant (Service, 2006). For this reason, most newer plants, especially outside the Middle East, use reverse osmosis (Jones, 2019).

1.2 Extractive Desalination

The present desalination technique (Chakrabarti, 2022) centers on a solvent-based liquid-liquid extraction method that exploits the solubility characteristics of a particular class of polar organic solvents. In particular, this method exploits the solubility of water, with an inverse temperature–solubility relation, in these solvents. In addition, like all organic solvents, this particular class of organic solvents does not dissolve salt (NaCl). The appropriate solvent is characterized by the following properties:

- dissolves water at or above 10% w/w at room temperature,
- has a boiling point above that of water, preferably above $150 \text{ }^\circ\text{C}$ under atmospheric pressure,
- has decreasing water solubility with increasing temperature (counter to typical solubility

behavior),

- is a liquid at room temperature,
- has differential solubility characteristics in saltwater and freshwater, such that the cloud point (the temperature at which two phases separate) is lower in saltwater than in freshwater, with the two cloud points being within about 20–100°C.

The last property is the key to the extraction methodology. At low temperature (i.e., ambient temperature) solubility of water in the solvent is high. The temperature is raised to the first cloud point and then to the second cloud point of the solvent. The different solubility of water in the solvent at the two cloud points permits separation of water from salts, as described next. The modest temperature range between the first cloud point and the second cloud point is important for cost-effectiveness.

The process is as follows. Starting with a solvent–saltwater mixture (Figure 1, A), desalination is carried out by phase separation in two steps, followed by an optional purification step.

1. The temperature of the solvent–saltwater mixture is raised to just above the solvent’s cloud point in saltwater. Just beyond the cloud point, the mixture separates into two clear phases: a freshwater-rich solvent phase (B) and a denser saltwater (brine) phase (C). During this phase separation, freshwater is extracted by the solvent and the water phase becomes salt-rich. The upper phase is processed in Step 2, and the lower phase is removed and discarded. Prior to being discarded, the brine phase can be passed through a biological treatment pond to biodegrade any trace residual solvent.
2. The freshwater-rich solvent phase (B) is heated to the solvent’s second cloud point, for freshwater, causing it to separate into two new phases: a solvent phase and a denser freshwater phase. After separation is complete, the new upper phase (D), now containing practically all the solvent, is recovered for reuse in Step 1. The freshwater lower phase (E) can be used directly (e.g., for agriculture or industry) or purified to potable levels in Step 3. Note that because solvent loss is extremely low, the bulk of the solvent can be reused repeatedly.

3. The freshwater phase (E) may be purified using conventional adsorption methods for removing polar organic solvents (e.g., activated carbon or bentonite clay), yielding water meeting drinking water standards. The adsorption media (activated carbon, clay, etc.) can be reused repeatedly, as in standard industrial purification. The trace amounts of solvent that are adsorbed would be disposed of in a customary industrial manner after the adsorption media loses its efficiency.

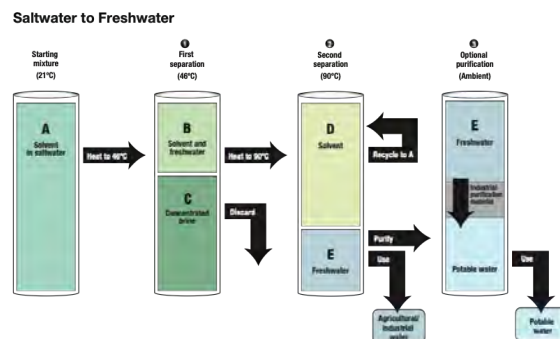


Figure 1. Schematic of novel extraction-based desalination with a dual-solubility polar organic solvent as reported here. (Proportions not to scale.)

2. Methods

2.1 Solvent Selection and Experimental Design

Candidate solvents were identified by searching the literature for compounds with appropriate characteristics. Screening focused first on polar solvents with low cloud points. Results of initial testing narrowed the focus to polar aliphatic compounds (either cyclic or acyclic). Because solubility data for some candidate compounds were unavailable or outdated, solubility curves were reconstructed or confirmed.

Promising candidates were tested to confirm the needed solubility characteristics. Initial tests broadly explored the efficiency of salt extraction (Step 1) and the extraction kinetics. In this initial screening, three candidates were tested: NP-9 (Dow Chemical), 2-[4-(2,4,4-trimethylpentan-2-yl)phenoxy]ethanol (Triton X-100, Dow Chemical), and *N*-cyclohexyl pyrrolidone (CHP, Dow Chemical). CHP had the best efficiency, and all remaining tests focused on this solvent. Next, small-volume tests with CHP were

done for several separation times to further test kinetics. Finally, both extraction steps were performed with CHP to obtain the final purity.

2.2 Measurement Methods

For all tests reported here, the solvent was mixed with a salt solution of approximately 3.5% salinity. In all cases, the solvent–brine mixture was initially held at 21°C in a constant-temperature water bath (1-chamber, 3-liter Digital Thermostatic Water Bath, Joan Lab Equipment Co., Ltd, Huzhou City, People’s Republic of China).

Separation of water from brine (Step 1) was achieved by heating the solvent–brine mixture to an appropriate temperature in the water bath (46°C for CHP) and holding at that temperature for the duration of the separation. After the separation period, the aqueous (lower, brine) phase was removed and measured for salinity using a water quality meter (3-in-1 digital pH meter with pH/total dissolved solids/temperature, unbranded). A calibration curve was generated (Figure 2) and used to determine salinity values.

Separation of freshwater from solvent (Step 2) was achieved by heating the first solvent (upper) phase from Step 1 to 90°C. The resulting second aqueous (lower, freshwater) phase was removed, and the volume was recorded to 2 significant figures. The salinity of the second aqueous phase was calculated by mass balance using the amount of NaCl present in the brine phase (to 3–4 significant figures; GEMINI-20 Portable Precision Digital Milligram Scale), and its solvent concentration was calculated using the solvent solubility curve.

Desalination efficiency was calculated by first drying the aqueous (brine) phase from the first separation step and weighing the amount of dry NaCl present. Then the percent desalination was calculated as weight of NaCl in aqueous phase divided by the calculated weight of starting NaCl (NaCl wt% x volume).

Purification of the freshwater (optional Step 3) was demonstrated with activated charcoal and bentonite clay. A solution was created that mimicked the <0.1% final CHP concentration in several of the extraction tests. The solution was agitated for 2

minutes with the charcoal or clay and then vacuumed twice through a Büchner funnel/filter setup. The amount of solvent was measured by boiling off the remaining water and weighing the excess material to 2 significant figures.

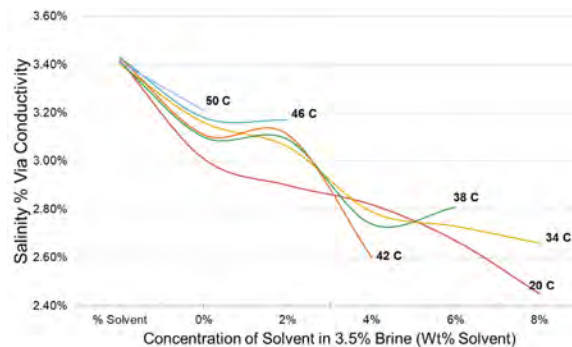


Figure 2. Salinity meter calibration curve, giving salinity vs. temperature and solvent concentration for *N*-cyclohexyl pyrrolidone (CHP, Dow Chemical).

3. Results

Solubility curves were recreated or confirmed for the candidate solvents. The curve for CHP is given in Figure 3.

In a screening test of desalination efficiency and kinetics, CHP far outperformed two other common nonionic surfactants, NP-9 and Triton X-100. Over a separation time of 3 hours, CHP reached 100% extraction (20% solvent; 46°C). In contrast, NP-9 had reached only about 60% extraction at 3 hours (5%, 60°C), and Triton X-100 showed no extraction at all (5%, 68°C). The test with Triton X-100 was allowed to run to completion: it took about 24 hours to reach 100% extraction. (Concentrations and temperatures for these tests were selected on the basis of solvent properties.)

Initial small-volume testing of the first extraction step showed that CHP had a calculated desalination efficiency of near 100% for all separation times tested (Table 1). Testing of the full process with a larger volume yielded a calculated desalination efficiency over 99% for a separation time of 3 hours (Table 2).

Purification tests with activated charcoal and bentonite clay confirmed that conventional purification methods yielded highly pure freshwater.

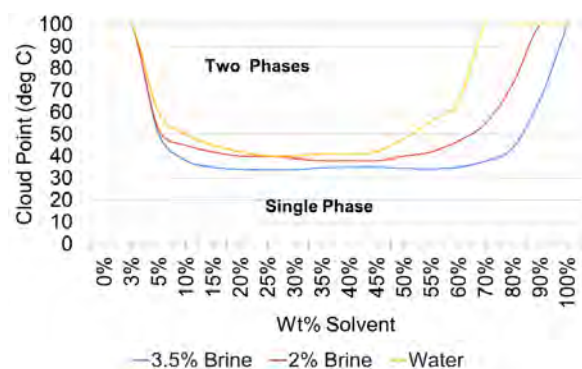


Figure 3. Solubility curve of *N*-cyclohexyl pyrrolidone (CHP, Dow Chemical) in 3.5% brine, 2% brine, and water. Solubility of the solvent is indicated by its cloud point: the temperature at which a solution (single phase) separates into water and solvent (two phases). The cloud point varies with the amount of solvent in the solution; this variation is shown in the shape of the solubility curve. The cloud point is also affected by the salt concentration of the solution, as shown by the different curves for 3.5% brine, 2% brine, and water (0% brine).

Table 1. Effect of first extraction duration on desalination efficiency with CHP.^a

Solvent Volume (mL)	Brine Volume (mL)	Solvent Load (%)	Salinity, Initial Brine (%)	Separation Time, Step 1, 46°C, (hours)	Salinity, Extracted Brine (%)	Calculated Desalination Efficiency (%)
2.8	11.2	20	3.51%	1	3.96	98.9
2.8	11.2	20	3.51%	2	4.18	100
2.8	11.2	20	3.51%	3	4.17	100

^aSolution held at 21°C in constant-temperature water bath until extraction protocol launched.

Table 2. Freshwater extraction from brine with CHP.^a

Solvent Vol. (mL)	Brine Vol. (mL)	Initial Brine Conc. (%)	Total NaCl (g)	Separation		Separation			Desalination Efficiency (%)	
				Time, Step 1, 46°C (h)	Adjusted Salinity, Extracted Brine (%)	Time, Step 2, 90°C (h)	Total NaCl in Brine Phase (g)	Total NaCl in Fresh-water (mg)		Volume Efficiency (%) ^b
10	40	3.51	1.40	1	4.08	1.40	2	1.87	11	99.0
20	80	3.51	2.81	1	3.97	2.78	2	29.18	10	91.7
20	80	3.51	2.81	2	4.04	2.79	2	15.00	11	96.0
20	80	3.51	2.81	3	4.07	2.81	2	1.49	11	99.6

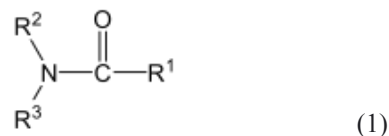
^a Solution held at 21°C in constant-temperature water bath until extraction protocol launched.

^b Volume efficiency calculated as volume of freshwater recovered per unit of input brine (i.e., units of freshwater per unit of saltwater).

4. Discussion

4.1 Surfactant Chemistry

Both aromatic or aliphatic solvents can be suitable, but aliphatic structures (either cyclic or acyclic) are preferred for biodegradability considerations. Structures include amides (Eq. 1), sulfones and sulfoxides (Eq. 2), and alcohols (Eq. 3):



where suitable structures of the R groups are defined in Chakrabarti (2022). These R groups are used to balance the overall polarity of the molecule and thereby determine the total solubility of the molecule in water.

To be usable for desalination using this method, the solvent must, in addition to the physical properties noted in section 1, derive its polarity from one or more groups selected from amide, substituted amide, sulfoxide, sulfone, hydroxyl, and ether groups.

Optimal candidates determined to date are shown in Table 3. Note that although the structure options specified in Chakrabarti (2022) serve to define an envelope of suitable structures, actual solubility properties and extraction efficiency must be verified by testing.

Table 3. Candidate polar organic solvents for low-temperature solvent-extraction desalination.

Compound
cyclohexyl pyrrolidone <i>also known as</i> <i>N</i> -cyclohexyl-2-pyrrolidinone <i>N</i> -cyclohexyl pyrrolidone
<i>N</i> -dodecyl pyrrolidone ethylene glycol monobenzyl ether 2-ethyl-butyl glycol <i>N</i> -hexyl glycol 4-methoxy-4-methyl-penta <i>N</i> -2-ol <i>also known as</i> 4-methoxy-4-methyl-pentanol-2

4.2 Benefits of This Extractive Approach

The properties of the solvents were chosen with a view to making the process practical for the widest possible range of applications and communities. The following were among the considerations.

The high extraction efficiency demonstrated for this class of solvent is important for three reasons. First, the process directly and relatively rapidly produces agriculturally usable water, without further treatment. Second, the very low concentration of solvent left in the freshwater fraction means that well-established industrial purification methods can be used to reach the purity required by regulatory bodies for drinking water (e.g., in the United States, the limit for foaming agents, i.e., surfactants, is 0.5 mg/L (U.S. Environmental Protection Agency, 2022). Purification tests confirmed complete or near-complete removal of the solvent, indicating that solvent concentration in treated water would be far

below the regulatory limit. Third, the completeness of the solvent–freshwater separation means the solvent can be reused many times before needing to be replaced, for lower operating costs.

In general, the process makes desalination accessible to communities across a broader range of geographic and economic conditions. Overall capital costs will be low because the process will use standard industrial extraction equipment. The relatively low solubility temperature limits the energy requirements for the process compared to distillation. The energy requirements are also low compared to reverse osmosis because no pressurization is required. As a result, this method would be feasible for communities whose access to fuel is limited by economics or geography (e.g., islands or remote coastal areas). Together, low-energy operation and closed-loop solvent design mean that a desalination option is now available for ecologically sensitive regions. Finally, the process will work for water with low initial salinity (brackish water); this characteristic, along with the low ecological impact, means that desalination becomes an option in relatively sensitive tidal river ecosystems.

Separate from municipal water supply and agricultural needs, a simple extraction method would also have advantages for marine applications where weight and energy consumption preclude use of existing technology.

5. Conclusion

The present work points the way to low-cost, energy-efficient desalination at industrial scale. Because extraction processes are ubiquitous, equipment design should not be a significant barrier to scale-up. However, because the process depends on solubility behavior with temperature, it will be important to explore the sensitivity of the extraction to temperature control. For designing for high throughput, the fast kinetics of CHP are promising: because extraction reaches 100% in 3 hours, there is room for optimizing dwell time vs. yield. For industrial scale-up, development will be required to optimize solvent recycling for long-term use. For example, potential regeneration cycles and replacement intervals will need to be investigated.

Predicting future water supply and demand on a global level is difficult because of the complex interactions among population growth, trade, industrialization, politics, climate change-related disruptions in precipitation, and improvements in conservation and efficiency (Oki and Kanae, 2006). Nevertheless, at the community and regional level, expanding the availability of usable water will open opportunities for economic development, improved quality of life, and greater resilience in areas that currently operate within a narrow and vulnerable window of viability.

Acknowledgments

I thank my father, Debtosh Chakrabarti, for encouraging me to look beyond what he achieved in his undergraduate research at MIT (Chakrabarti, 1997). I thank Ms. Jane Marie Andrew for helping me in properly structuring the manuscript. I also thank my science teachers, including Mr. Sean Watson, for taking an active interest in the research project.

References

- Bierkens, M. F. P., & Wada, Y. (2019). Non-renewable groundwater use and groundwater depletion: A review. *Environmental Research Letters*, 14(6), 063002. <https://iopscience.iop.org/article/10.1088/1748-9326/ab1a5f>
- Chakrabarti, D. (1997, October 21). Desalination of seawater by nonionic surfactant aided phase separation. U.S. Patent 5,679,254.
- Chakrabarti, D. (2022, June 3). Extractive desalination of sea water using a special class of polar organic solvents. U.S. Patent application serial number 17/805,321.
- Chibani, A. (2022, July 14). *From the river to the sea: Water management in Libya*. The Tahir Institute for Middle East Policy. <https://timep.org/commentary/analysis/from-the-river-to-the-sea-water-management-in-libya/>, accessed October 7, 2022.
- Food and Agricultural Organization (FAO) of the

United Nations (2020). *AQUASTAT—FAO's global information system on water and agriculture*. <http://www.fao.org/nr/water/aquastat/main/index.stm>

Gasson, C. (2020, May 28). *We have \$0.40 desal. Where do we go from here?* Global Water Intelligence. <https://www.globalwaterintel.com/news/2020/22/we-have-0-40-desal-where-do-we-go-from-here>, accessed October 7, 2022.

Gleick, P. H., & Cooley, H. (2021). Freshwater scarcity. *Annual Review of Environment and Resources*, 46, 319–348. <http://doi.org/10.1146/annurev-environ-012220-101319>

Hofste, R. W., Reig, P., & Schleifer, L. (2019, August 6). *17 countries, home to one-quarter of the world's population, face extremely high water stress*. World Resources Institute. <https://www.wri.org/insights/17-countries-home-one-quarter-worlds-population-face-extremely-high-water-stress>, accessed September 28, 2022.

Jones, E., et al. (2019). The state of desalination and brine production: A global outlook. *Science of the Total Environment*, 657, 1343–1356. <http://doi.org/10.1016/j.scitotenv.2018.12.076>

LiVecchi, A., et al. (2019, April). *Powering the blue economy: Exploring opportunities for marine renewable energy in maritime markets*, Chapter 7, Desalination. U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy. <https://www.energy.gov/sites/prod/files/2019/03/f61/73355.pdf>

Oki, T., & Kanae, S. (2006). Global hydrological cycles and world water resources. *Science*, 313(5790), 1068–1072. <http://doi.org/10.1126/science.1128845>

PopulationU. (n.d.). *Countries by GDP*. <https://www.populationu.com/gen/countries-by-gdp>, accessed October 7, 2022.

Service, R. F. (2006). Desalination freshens up. *Science*, 313(5790), 1088–1090. <http://doi.org/10.1126/science.313.5790.1088>

Trenberth, K. E., Fasullo, J. T., & Mackaro, J. (2011). Atmospheric moisture transports from ocean to land

and global energy flows in reanalyses. *Journal of Climate*, 24, 4907–24, Figure 9, <http://doi.org/10.1175/2011JCLI4171.1>

U.S. Environmental Protection Agency (EPA). (2022, February 17). *Secondary drinking water standards:*

Guidance for nuisance chemicals. Table of secondary standards.

<https://www.epa.gov/sdwa/secondary-drinking-water-standards-guidance-nuisance-chemicals>, accessed October 6, 2022.

Understanding the Relationship Between Religiosity, Religious Affiliation, Prior Knowledge of Nanotechnology, and the Ethical Attitudes of Nanotechnology

Rumaysa Ahmad^{1*}

¹Granada Hills Charter High School, Granada Hills, CA, USA

Received August 14, 2022; Revised February 12, 2023; Accepted, March 10, 2023

Abstract

Nanotechnology has been a critical field of study through its integration of many appliances. With public trust conflicting with the acceptance of nanotechnology, researchers have established that certain variables - such as prior knowledge of nanotechnology, religion, and religiosity - play a significant role in shaping people's ethical perceptions of nanotechnology and, eventually, their public opinions. Therefore, this study focuses on finding and building the relationships between knowledge, religion, religiosity, and ethical concern of nanotechnology among teenagers in Granada Hills, California, to understand the influence of these variables on student perceptions of nanotechnology. A correlational study using quantitative data including a demographic survey, a 15-question Centrality of Religiosity Scale, prior knowledge evaluation, and an ethical concern section was administered to high school students from two schools in Granada Hills. Pearson product-moment correlations, ANOVA, and descriptive statistics were used to analyze the data. The results depicted that there was a negative relationship between knowledge and mean ethical concern, a negative relationship between knowledge and mean religiosity, a positive relationship between mean religiosity and mean ethical concern, and a positive correlation between age and mean knowledge of nanotechnology, all with weak coefficients. Ethical concerns were also explored, and students were found to be the most concerned about nanotechnology getting in the "wrong hands" and the loss of freedom and privacy. Additional research should be conducted on greater populations that consider different sociodemographic, affective, sociodemographic, or cognitive variables.

Keywords: Nanotechnology, Ethical Concern, Public Trust, Religiosity, Religion, Knowledge

1. Introduction

Throughout the recent decade, nanotechnology has made an increasingly relevant presence among society and many sectors of public well-being. Whether it is through medical applications, transportation, food, or even environmental conservation, the versatile nature of nanotechnology has been manipulated and integrated to improve

many foundational sectors. According to physicist Richard Feynman, the scientific definition for nanotechnology is the application of atoms at the nanoscale across various fields including chemistry, biology, medicine, and engineering (NNI). With greater advancements, nanotechnology has the capability to improve the human condition through efficient disease prevention, sustainable practices, and water filtration for poverty-stricken communities.

* Corresponding Author
rumaysa16.mengo@gmail.com

Advisor: Dr. Spencer Wolf
swolf@ghctk12.com

For instance, nanotechnology has significantly aided cancer treatments because of its increased “biocompatibility, reduced toxicity, more excellent stability, enhanced permeability and retention effect, and precise targeting” (Gavas et al., 2021). Therefore, there is no doubt that nanotechnology will be the future of science and be used to target global crises. However, despite the market value for nanotechnology rising to an astounding \$76 billion, the lack of public trust among consumers has only hampered national acceptance of nanotechnology. Among the first polls conducted in 2004, the National Science Foundation found that 49% of the respondents had “heard nothing at all” about nanotechnology, while many believed preconceived notions about its practical use, moralities, and accessibilities (Institute of Medicine (US) Food Forum). Four years later, a similar poll was conducted, only to find that an increasingly large portion of people still had limited knowledge about nanotechnology and failed to give their thoughts about its risks and benefits. Since Americans have yet to gain public awareness of nanotechnology, commercialization of modernized technologies has progressed slower (Peter D. Hart Research, Inc. 2008). Julia Moore, director of the Wilson Center, argued that misinformation or a lack of knowledge correlates with unguided public approval for nanotechnology, stating that “public opinion is really up for grabs when it comes to nanotechnology” (Institute of Medicine (US) Food Forum). Among the many other reasons for distrusting nanotechnology, religiosity, defined as the “behavior and beliefs associated with organized religion” is also becoming an important variable, as previous studies have demonstrated that being more religious meant having less trust in science (Good et al., 2011, pg 538). According to a survey incorporating 1,015 participants done by the Center for Nanotechnology, Americans had shown a strong negative correlation between religiosity and the acceptance of nanotechnology (Scheufele et al. 2009). Finally, acknowledging the ethical dilemmas that people may have about nanotechnologies, such as its military uses, environmental consequences, and medical complications, may allow the concerns of consumers to be heard to prevent widespread backlash or

skepticism (Macoubrie, 2006). Therefore, this paper strives to explore how religious alliances, prior knowledge, and ethical issues may correlate with the limited social approval of nanotechnologies among consumers.

1.1 Theology of Religion and Religiosity Within Nanotechnology

Religion has continually had a lasting impact on people’s opinions regarding technology and its vast outreach. Franz Foltz, Assistant Professor at Rochester Institute of Technology, and Fredrich Foltz specifically examined the Christian response to nanotechnology by describing the four naive views of religion. One of the views they mentioned was that certain morals such as human dignity or social justice might be amplified with nanotechnology. For instance, biotechnology used to manipulate human characteristics is “not preserving human dignity” since the Roman Church believes this is playing the hands of God (F. Foltz and FR. Foltz, 2006). Similarly, Kotze, Manitza (2018) attempted to look at the bioethical issues that genetic engineering and robotics raised from a theological perspective. Transhumanism, a word only popularized in the 1950s, meaning the enhancement of human beings and what it means to be human has posed conflicts for religious individuals’ perceptions of nanotechnology. The World Transhumanist Association released a statement stating that on one hand, bioelectronics enables humans to perform at a high- functioning level, while on the other hand, the human body is almost “dematerialized” into “a partially artificial being- a cyborg” (Kotze, 2018, pg. 1). She also discussed how disapproval for transhumanism could reflect onto people’s public trust, public opinion, and medical science views on nanotechnology.

Other than religious affiliation, religiosity has already proven to be a determining factor in people’s disposition for nanotechnology (Good et al., 2011, pg 538). National data was used to examine how people utilized science media, knowledge about nanotechnology, and predispositions, such as the strength of religious beliefs, to form their attitudes about nanotechnology. A national telephone survey

of 706 respondents found that there seemed to be a significant negative relationship between the strength of religious beliefs and support for funding (Brossard et al., 2009). Agreeing with Franz Foltz and Fredrich Foltz, people with a strong connection to religion had been opposed to funding nanotechnology because of their belief that it had human-altering properties (Brossard et al., 2009). While understanding ethical concerns, Scheufele et al. (2009) surveyed more than 1,015 Americans and 29,193 Europeans and found that there was a negative correlation between religiosity and the morality of technology. They identified that strong religiosity often led to citizens perceiving nanotechnology as ethically wrong. Among European countries, there were more positive outlooks on nanotechnology with a weaker connection to religiosity, which may be explained by the secularized education system of many European nations (Scheufele et al. 2009). Therefore, being highly religious, or having a strong connection with religion and participating in religious practices, has been shown to have a negative correlation with the general ethical perceptions of nanotechnology.

However, most research studies took an approach looking at older populations, ranging from 25–60-year-olds, with less focus on the youth. According to data released by the PEW Research Center, about 50% of teens responded that they either share the same religious values as their parents or differ slightly. With the recent uprise in the youth identifying as Atheist or Agnostic, 66% of teens who do religious practices with their family say they do so partly because their parents want them to (Diamant and Sciupac, 2020). Since a greater societal role is being played by teens, many ideologies about nanotechnology, passed down by parents, may affect the public opinions that teens hold on consumer products. Additionally, identifying the correlation that younger generations have with religion and religiosity may be the key to understanding their opinions about the uses of nanotechnology in order to better adapt to their concerns.

1.2 Knowledge of Nanotechnology

Prior knowledge of nanotechnology also has the potential to influence public support and funding for

them. In the early 2000s, there was an influx of science-fiction culture and optimism for nanotechnology, even though there was a lack of knowledge of what nanotechnology was. Although the percentage of people who have “heard nothing at all” about nanotechnology decreased from 2006 to 2010, more than half of the population has limited knowledge of nanotechnology (Binder, 2013). Dominique Brossard and her peers at the University of Wisconsin-Madison identified that religiosity is also connected with knowledge, in that highly religious individuals were shown to have lower knowledge of nanotechnology and therefore, had lower reasons to support it (Brossard et al., 2009). Their misconceived views could have not only been reinforced by their strong religious connection but could also have been passed down to their children. Jane Macoubrie’s study (2006) utilized a quasi-experimental group and established that public policies and concerns about nanotechnology must be addressed by people who are knowledgeable about nanotechnology for it to be properly evaluated.

In the study presented by Gardner, Grant et al. (2010), the “knowledge deficit model” claims that increased knowledge about a complex topic “promotes more rational formation of attitudes” (Gardner et al., 2010). Hence, his research question of understanding the risk perceptions of nanotechnology came from questioning undergraduate students in the engineering fields. Gardner explains that since undergraduate courses often are the first contact that students have with nanotechnology discourse, this is where their attitudes start forming about the practices of nanotechnology. Therefore, undergraduates exposed to nanoscale science were chosen for his study because of their greater education in nanotechnology. Both studies reflect that through greater education and knowledge of nanotechnology, people may be able to bridge the gap between skepticism and national acceptance of its integration. Furthermore, to generate effective dialogue over the values, visions, and societal implications of nanotechnology, an effective ‘nanotechnology engagement project’ should be developed. One that uses a two-way dialogue system between scientists and the public will impact regulations of nanotechnology (Pidgeon

et al., 2011).

1.3 Theoretical Framework

Since public trust remains a relevant issue in gaining approval for technologies, such as nanotechnology, it is vital to discuss how to address the concerns of laypeople (non-experts). According to researcher Trond Am, from the Norwegian University of Science and Technology, there are two key reasons for the lack of concern that people have expressed against technological advancements: 1) prior technological controversies that have been poorly handled by the government and 2) the ambiguity within scientific advances that run ahead of public awareness and control (Am, 2011). Therefore, tools such as the Nanotechnology-Perception Attitude Acceptance Framework (Nano-PAAF), developed by Dr. Achintya Bezbaruah, and Dr. Rajesh Pillai (2017), help build a systemic understanding of the phenomenon by addressing how certain consumer behaviors impact public trust and attitudes toward nanotechnology. The framework proposed that people's perceptions of nanotechnology are influenced by cognitive, affective, socio-cultural, and sociodemographic factors. Among the cognitive factors, Pillai and Bezbaruah reinforced the idea that knowledge about nanotechnology has been related to greater support for it and a positive attitude toward science (Scheufele and Lewenstein, 2005). Compared to experts, laypeople perceive greater risks and lower support for new technologies, which is why it may be useful to identify their prior knowledge to get a better indicator of where their attitudes lie. Additionally, sociocultural factors such as a strong religious standing was correlated with lower levels of trust within nanotechnology. Nano-PAAF suggested that "non-experts are more likely to use religious beliefs as a heuristic cue to assess nanotechnology risks as compared to experts" even when factors like trust, knowledge, and media exposure are controlled (Pillai and Bezbaruah, 2017, pg. 7). Overall, identifying the factors expressed in the Nano-PAAF can be a good indicator of the ethical concerns that consumers have toward nanotechnology based on their perceived risks and benefits. Eventually, these concerns can be

efficiently tackled through regulations and public policies. Hence, Nano-PAAF is used to observe how prior knowledge and religiosity may shape the ethical attitudes consumers have toward nanotechnology.

1.4 Research Question

After analyzing past research related to the ethical perceptions of nanotechnology and the variables suggested before, there has been a lack of research correlating all four of the variables: religious affiliation, religiosity, prior knowledge, and ethical perceptions. More specifically, previous studies did not take any approach to understand the role that these variables played among teenagers. Therefore, this research paper will try to bridge this gap through an exploratory study of the relationship between these variables.

In addition, the Granada Hills region in California was analyzed as a starting point to guide future research on greater regions in the US. Many schools in Granada Hills are diverse with approximately 75% minority enrollment and a range of students from various socioeconomic backgrounds. Therefore, Granada Hills was chosen as an ideal location to gain perspectives from students of different upbringings and backgrounds, leading to the research question: what is the relationship between religious affiliation, religiosity, prior knowledge of nanotechnology, and the ethical concerns that high school students in the Granada Hills region have on nanotechnology?

Exploring this research question will allow public policymakers and nanotechnology innovators to cater to and address the ethical concerns that teenagers may have, considering their varying religious and socio-cultural backgrounds. Moreover, scientists can gain a deeper insight into the vulnerabilities of public trust through the lens of teenagers, who will eventually gain control over the industry.

2. Methods

This study used a quantitative survey to explore the correlation between religious affiliation, religiosity, prior knowledge of nanotechnology, and the ethical concerns that students have. This method was used to not only observe the relationships

between the variables but to do a correlational analysis of the various religions and ethical concerns using quantitative data.

2.1 Population

Data was collected from high schoolers in grades 9 -12 mainly because prior data concerning ethical attitudes toward nanotechnology failed to consider the perceptions of teens. Regardless, the teenager perspective is beneficial since they will soon serve as leading innovators and regulators of new technology, including nanotechnology. Chris Toumey, a researcher from the University of South Carolina, expressed how public reactions to nanotechnology must come from both scientific experts and regular citizens, specifically teens, and their perspectives (2004). Two major high schools in the Granada Hills region, referred to as School 1 and School 2, were analyzed. Both samples have the greatest student population compared to other high schools in the region and incorporate students from various religions: Christians (under the Protestant, Catholic, and Orthodox sectors), Atheists, Agnostics, Muslims, Hindus, and Jews. Granada Hills, California, was chosen because of its diverse religious population and progressive push in STEM curriculum (Patten and Newhart, 2018, pg. 52). Studying this population allowed the gap within my research question to be addressed, by analyzing whether the correlation between knowledge, religiosity, religion, and ethical concerns among teens are similar to the data collected from adults or whether there are overwhelming differences that require a different approach. Since it has been observed that teens are slowly becoming “disconnected” from religion and loosely following their guardians, it was crucial to observe whether religion and religiosity remained principal factors in determining attitudes toward nanotechnology as Nano-PAAF had suggested, or if there was a need to evolve the theory based on the newer generations (Good et al. 2011).

2.2 Instrument

The selected instrument was a survey sent through a Google Form that had 4 major sections which

correlated with the points of interest. Primarily, there was a demographic questionnaire that asked the participant’s name, grade level, age, school, race, and religious group, which also included an open-ended portion if they needed to elaborate on their current religious beliefs. This section was essential in recognizing other variables outside of my research scope that could influence the variables. The next section included a replica of the 15-question Centrality of Religiosity Scale (CRS) which is used to measure how religious a person is based on five core dimensions of religiosity: public practice, private practice, religious experience, ideology, and religious intelligence (S. Huber and O. Huber, 2012). The scale has been used across 25 different countries and has even been validated to support cross-cultural barriers and non-Abrahamic religions. Each item is scored on a scale of 1 to 5 and then the sum score is divided by the number of scored scale items. If the resulting number is between 1.0 to 2.0, they are not religious, 2.1 to 3.9 means they are religious, and 4.0 to 5.0 means that they are highly religious (S. Huber and O. Huber, 2012).

The following section was a simple Likert Scale adapted from Joubert et al. and their study on the current levels of knowledge and attitudes toward nanotechnology among the Austrian population. It asked, “To what extent do you feel informed about nanotechnology?” with scores ranging from 1 (I don't know anything about nanotechnology) to 4 (I am well versed in nanotechnology) (2020). Finally, the last section started with a short information page describing nanotechnology and its application to provide background information to the respondents before proceeding. That page included information quoted directly from the National Nanotechnology Initiative, Stephen J Florczyk and Subrata Saha's research paper, "Ethical issues in Nanotechnology," and "Applications of Nanotechnology in Daily Life" by Mahmoud Nasrollahzadeh et al. Jane Macoubrie, adjunct Professor at the University of Southern California, did a similar study where participants were given information cards about nanotechnology, then told to formulate potential ethical concerns that concerned them (Macoubrie, 2006). In the same manner, a Likert scale was created for each of those ethical concerns found in her study and asked the

respondents to rate them on a scale of 1 (not at all concerning) to 5 (very concerning). It was not left as an open-ended question because a limitation in my pilot study was that many students could not think of ethical concerns from the top of their heads.

2.3 Recruiting Method

To select students to take the survey, a systematic sampling method was utilized by contacting every 5th teacher on the school’s faculty page to distribute the instrument to their classes. To reduce bias, teachers from every subject were chosen and there was no preference for male or female teachers (Patten and Newhart, 2018). Teachers were also listed alphabetically by each subject, however, only those teaching physical education or ESL were not contacted due to the inaccessibility of the instrument adapting to their learning environment. The trends and data were analyzed from all the students who participated in the survey in order to generalize the teens in Granada Hills, excluding those who fell into the category of minority religions (Patten and Newhart, 2018). Probabilistic sampling was utilized to collect a representative sample and compare the religious differences in the sample rather than selecting specific members of each religious identity.

2.4 Data Collection

After sending out the survey to 32 teachers, I received confirmation that those teachers had posted the link to their main platform for their students to take, by choice. Google Forms was used to distribute the survey because it was user-friendly, did not support duplicate answers, could automatically be linked to Google Sheets, and was the most accessible to teens.

To ensure no possible ethical concerns, samples of convenience were not used. Instead, a diverse population including students from regular classes to AP-level classes who would have various levels of knowledge about nanotechnology was examined (Patten and Newhart, 2018). All the questions that came from previous scales were validated while the personal questions were approved by the International Review Board. In addition, before

taking the survey, participants were given a consent form, informing them of what their data was going to be used for, the type of questions asked, and how long the form would take, giving them the option to opt-out anytime.

3. Results

To test the relationships with the prior variables, I made a few adjustments to the sample size and collected demographic data as depicted in Table 1. The Jewish and Mormon religious groups had not reached a significant number of responses, so they were not utilized when doing any statistical tests.

Table 1: Descriptive statistics of the demographics of respondents

	Frequency (n)	Mean (\bar{x})	Standard Deviation (σ)
Total Respondents	162		
Age	-	15.8	±0.99
14	14	-	-
15	40	-	-
16	63	-	-
17	39	-	-
18	3	-	-
19	1	-	-
Mean Religiosity	-	2.93	±1.02
Mean Prior Knowledge	-	1.72	±0.71
Race			
White	46	-	-
Black/African American	14	-	-
Native American and Alaska Native	2	-	-
Asian	64	-	-
Native Hawaiian and Other Pacific Islander	7	-	-
Hispanic	48	-	-
Other	17	-	-
Religion			
Agnostic	18	-	-
Atheist	28	-	-
Catholic	60	-	-
Jewish	2	-	-
Mormon	2	-	-
Muslim	7	-	-

Non-Abrahamic	8	-	-
Orthodox Christian	19	-	-
Protestant Christian	15	-	-

of 15 and 16, or 10th and 11th graders. The religious groups that were used in the data analysis process were Agnostic, Atheist, Catholic Christian, Muslim, Non-Abrahamic (including Hinduism and Buddhism), Orthodox Christian, and Protestant Christian, which is representative of the Granada Hills religious breakdown (Loksata, 2012)

Based on the demographic data, most students who completed the survey were roughly between the ages

3.1 Overall Ethical Concern

Table 2. Descriptive statistics of the level of concern among the following ethical issues in nanotechnology

	Mean Ethical Concern	Long-Term Health	Military Uses	Environmental Footprint	Social Footprint	Lost Freedoms
Response Number	161	161	161	161	161	161
Mode	3.000	3.000	4.000	4.000	3.000	5.000
Median	3.636	3.000	4.000	3.000	3.000	4.000
Mean	3.501	3.068	3.553	3.205	2.969	3.845
Std. Deviation	0.730	1.146	1.204	1.256	1.075	1.132

	Lack of Regulation	Losing Funding	Effects on Nature	Getting in “Wrong Hands”	Public Trust	Responsibly Researching
Response Number	161	161	161	161	161	161
Mode	3.000	3.000	4.000	5.000	3.000	3.000
Median	4.000	3.000	4.000	5.000	3.000	3.000
Mean	3.702	3.298	3.665	4.242	3.534	3.429
Std. Deviation	1.117	1.106	1.199	0.980	1.025	1.192

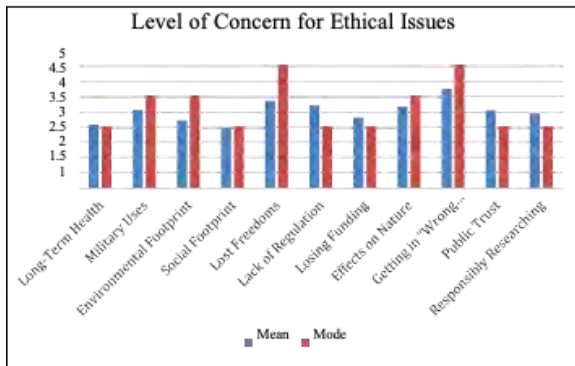


Figure 1. Bar graph depicting the mean and mode for each ethical concern on a scale of 1 (not at all concerning) to 5 (very concerning).

When analyzing the overall ethical concerns that the students had for nanotechnology, the following categories were used: long-term health concerns,

military uses, environmental footprint, social footprint, loss of freedoms and privacy, lack of regulation, losing funding for other priorities, effect on nature, getting in the “wrong hands,” public trust, and responsibly researching, adapted from Joubert et al., 2020 and the ethical concerns she collected from her study of nanotechnology. Among the concerns illustrated in Table 2 and illustrated in Figure 1, students were most concerned about nanotechnology getting into the “wrong hands” and the loss of freedoms and privacy that may develop from using nanotechnology. Overall, students seemed to have a neutral understanding of each concern with the lowest concern being the social footprint of nanotechnology, or the social implications.

A Pearson product-moment correlation coefficient was also calculated to see the relationships between each ethical concern. As shown in Table 3, most of

the concerns had shown a statistically significant positive correlation with each other, with the strongest correlations being environmental footprint vs. long-term health ($p < 0.001$, $r = 0.488$), environmental footprint vs. effects on nature ($p < 0.001$, $r = 0.580$), loss of freedoms vs. lack of regulation ($p < 0.001$, $r = 0.571$), getting in the “wrong hands” vs. public trust ($p < 0.001$, $r = 0.517$), responsibly researching vs. long-term health ($p < 0.001$, $r = 0.525$), and responsibly researching vs. public trust ($p < 0.001$, $r = 0.498$). Within each of these relationships, the data was significant since the p-value was less than 0.001, and one variable seemed to be strongly interconnected with another. For example, the more concerned students felt about the environmental footprint of nanotechnology, the more concerned they also felt about its long-term health effects. Therefore, many of the ethical concerns were strongly positively associated.

Table 3. Pearson’s correlation of all the ethical concerns

Pearson’s Correlations	Pearson's r	p
Environmental Footprint - Loss of Freedoms	0.203*	0.011
Environmental Footprint - Lack of Regulation	0.158*	0.047
Environmental Footprint - Losing Funding	0.207**	0.009
Environmental Footprint - Getting in Hands"	0.277**	<.001
Environmental Footprint - Responsibly Researching	0.483***	<.001
Environmental Footprint - Long-Term Health	0.488***	<.001
Environmental Footprint - Military Uses	0.312***	<.001
Environmental Footprint - Social Footprint	0.414***	<.001
Environmental Footprint - Effects on Nature	0.580***	<.001
Environmental Footprint - Public Trust	0.211**	0.008
Loss of Freedoms - Lack of Regulation	0.571***	<.001
Loss of Freedoms - Losing Funding	0.392***	<.001
Loss of Freedoms - Getting in "Wrong Hands"	0.450***	<.001
Loss of Freedoms - Responsibly Researching	0.366***	<.001

Loss of Freedoms - Long-Term Health	0.386***	<.001
Loss of Freedoms - Military Uses	0.247**	0.002
Loss of Freedoms - Social Footprint	0.278***	<.001
Loss of Freedoms - Effects on Nature	0.418***	<.001
Loss of Freedoms - Public Trust	0.377***	<.001
Lack of Regulation - Losing Funding	0.439***	<.001
Lack of Regulation - Getting in "Wrong Hands"	0.477***	<.001
Lack of Regulation - Responsibly Researching	0.345***	<.001
Lack of Regulation - Long-Term Health	0.336***	<.001
Lack of Regulation - Military Uses	0.298***	<.001
Lack of Regulation - Social Footprint	0.332***	<.001
Lack of Regulation - Effects on Nature	0.344***	<.001
Lack of Regulation - Public Trust	0.470***	<.001
Losing Funding - Getting in "Wrong Hands"	0.440***	<.001
Losing Funding - Responsibly Researching	0.407***	<.001
Losing Funding - Long-Term Health	0.305***	<.001
Losing Funding - Military Uses	0.121	0.130
Losing Funding - Social Footprint	0.252**	0.001
Losing Funding - Effects on Nature	0.384***	<.001
Losing Funding - Public Trust	0.388***	<.001
Getting in "Wrong Hands" - Responsibly Researching	0.461***	<.001
Getting in "Wrong Hands" - Long-Term Health	0.355***	<.001
Getting in "Wrong Hands" - Military Uses	0.348***	<.001
Getting in "Wrong Hands" - Social Footprint	0.298***	<.001
Getting in "Wrong Hands" - Effects on Nature	0.436***	<.001
Getting in "Wrong Hands" - Public Trust	0.517***	<.001
Responsibly Researching - Long-Term Health	0.525***	<.001

Responsibly Researching - Military Uses	0.196*	0.014
Responsibly Researching - Social Footprint	0.403***	<.001
Responsibly Researching - Effects on Nature	0.450***	<.001
Responsibly Researching - Public Trust	0.498**	<.001
Long-Term Health - Military Uses	0.228**	0.004
Long-Term Health - Social Footprint	0.328***	<.001
Long-Term Health - Effects on Nature	0.420***	<.001

*p < 0.05, **p < 0.01, ***p < 0.001

3.2 Religion vs. Religiosity

ANOVA was used to test the statistical significance between the mean religiosity scores of each religion. The test assumptions were checked, but Levene’s test was significant (p = 0.001) since the sample size for each religion varied. Therefore, the null hypothesis that each religion had statistically insignificant religiosity variances was rejected. Normality was checked with a Q-Q plot and no deviations were noticed, which means that the data was normally distributed (Table 4 and 5).

There was a significant difference among the 7 religions studied and their religiosity scores (p < 0.001). Since the p-value was < 0.001, there was stronger evidence for the difference in means or that the data was statistically significant. Post hoc testing revealed significant differences between the religiosity scores from the Agnostic (Mean = 2.2, SD = 0.44), Atheist (Mean = 1.6, SD = 0.53), and Non-Abrahamic (Mean = 2.5, SD = 0.59) religions compared to Catholic (Mean = 3.3, SD= 0.73), Muslim (Mean = 3.8, SD = 0.82), Orthodox Christian (Mean = 3.3, SD = 0.96), and Protestant Christian

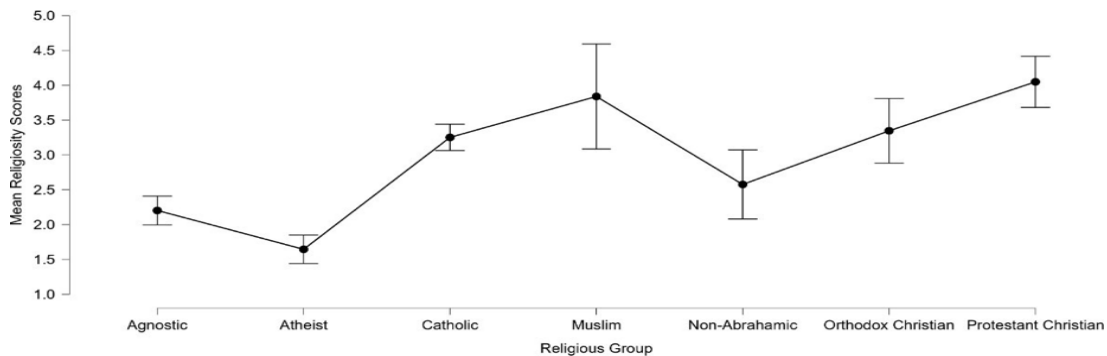


Figure 3. Descriptive plot showing the difference in religiosity scores among religious groups through ANOVA.

(Mean = 4.0, SD = 0.66) (Table 6). The data showed that there was a clear difference in religiosity scores among the religions as lower means depicted lower religiosity while higher scores meant the opposite (Table 7). Atheism was shown to be the least religious, while Protestant Christians were shown to be the most devoted to practicing religion including worship and personal prayer.

Table 4. ANOVA results measuring mean religiosity among religious groups

Cases	Sum of Squares	df	Mean Square	F	p
Religious Group	91.747	6	15.291	31.906	<.001
Residuals	71.888	150	0.479		

Table 5. Assumption check using Levene’s Test for Equality of Variances

F	df1	df2	p
3.981	6.000	150.000	0.001

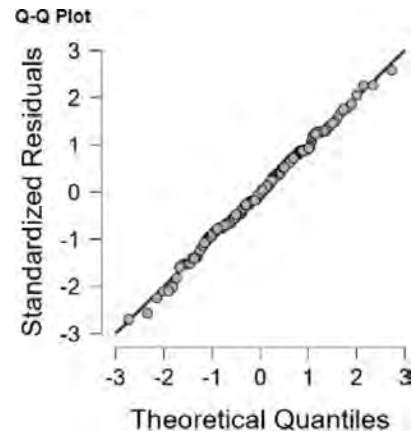


Figure 2. Q-Q plot is used to show that data points are normally distributed but are not equally varied through Levene’s test for religiosity scores.

Table 6. Post Hoc comparisons among religious groups

		Mean Difference	95% CI for Mean Difference		SE	t	P _{tukey}
			Lower	Upper			
Agnostic	Atheist	0.557	-0.048	1.163	0.203	2.750	0.093
	Catholic	-1.049	-1.583	-0.515	0.179	-5.858	<.001
	Muslim	-1.636	-2.544	-0.727	0.304	-5.380	<.001
	(Non-Abrahamic)	-0.373	-1.239	0.492	0.290	-1.289	0.856
	Orthodox Christian	-1.143	-1.806	-0.480	0.222	-5.153	<.001
	Protestant Christian	-1.846	-2.552	-1.139	0.236	-7.805	<.001
Atheist	Catholic	-1.606	-2.080	-1.133	0.158	-10.137	<.001
	Muslim	-2.193	-3.067	-1.319	0.293	-7.496	<.001
	(Non-Abrahamic)	-0.931	-1.760	-0.101	0.278	-3.353	0.017
	Orthodox Christian	-1.700	-2.315	-1.085	0.206	-8.262	<.001
	Protestant Christian	-2.403	-3.065	-1.741	0.222	-10.848	<.001
Catholic	Muslim	-0.587	-1.413	0.240	0.276	-2.122	0.345
	(Non-Abrahamic)	0.676	-0.103	1.454	0.261	2.593	0.136
	Orthodox Christian	-0.094	-0.639	0.451	0.182	-0.516	0.999
	Protestant Christian	-0.797	-1.394	-0.200	0.200	-3.987	0.002
Muslim	(Non-Abrahamic)	1.262	0.191	2.333	0.358	3.523	0.010
	Orthodox Christian	0.493	-0.422	1.408	0.306	1.610	0.676
	Protestant Christian	-0.210	-1.157	0.737	0.317	-0.663	0.994
(Non-Abrahamic)	Orthodox Christian	-0.770	-1.642	0.102	0.292	-2.637	0.122
	Protestant Christian	-1.472	-2.378	-0.567	0.303	-4.858	<.001
Orthodox Christian	Protestant Christian	-0.703	-1.418	0.012	0.239	-2.940	0.057

*P-value and confidence intervals adjusted for comparing a family of 7 estimates (confidence intervals corrected using the tukey method).

Table 7. Descriptive statistics of mean religiosity scores

Religious Group	Mean	SD	N
Agnostic	2.203	0.441	20
Atheist	1.646	0.528	28
Catholic	3.252	0.730	60
Muslim	3.839	0.815	7
Non-Abrahamic	2.579	0.594	8
Orthodox Christian	3.346	0.961	19
Protestant Christian	4.049	0.664	15

3.1 Religion vs. Prior Knowledge

To test the difference in prior knowledge among different religious groups, one-way ANOVA was used again. The test assumptions were checked, and Levene’s test was non-significant (p = 0.288), meaning that there was not enough variance in the data to account for the difference in means (Table 8).

However, there seemed to be no significant difference among the religions since the p-value = 0.965 and there was a low F-value = 0.234 indicating that there was not much variation in the actual samples (Table 10). According to the descriptive plot in Figure 4, prior knowledge of nanotechnology did not differ among religious groups, as each had a high standard deviation including the seemingly religious ones.

Table 8. Levene’s Test showing minimal variance in data for knowledge in different religious groups

F	df1	df2	p
1.242	6.000	150.000	0.288

Table 9. ANOVA results measuring mean knowledge scores among religious groups

Cases	Sum of Squares	df	Mean Square	F	p
Religious Group	0.743	6	0.124	0.234	0.965
Residuals	79.359	150	0.529		

Table 10. Descriptive statistics of respondents’ knowledge of nanotechnology

Religious Group	Mean	SD	N
Agnostic	1.600	0.821	20
Atheist	1.786	0.738	28
Catholic	1.733	0.756	60
Muslim	1.857	0.378	7
Non-Abrahamic	1.750	0.707	8
Orthodox Christian	1.632	0.597	19
Protestant Christian	1.667	0.724	15

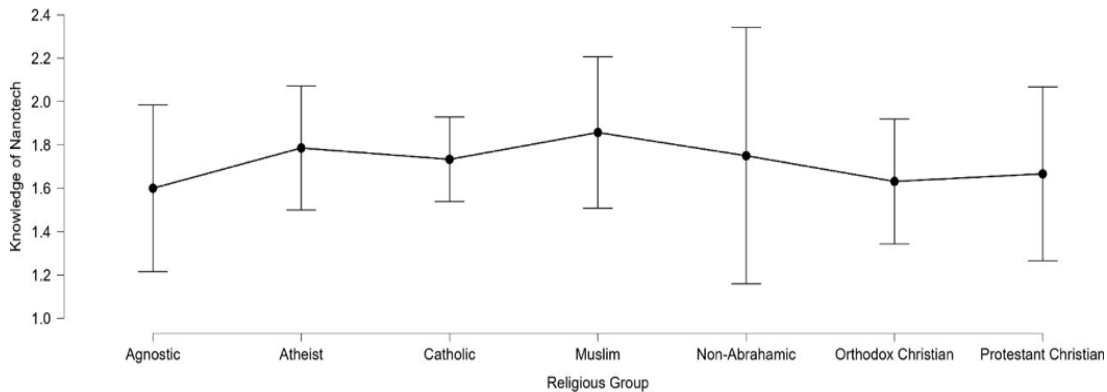


Figure 4. The graph depicts the comparisons of mean knowledge scores across the given religions.

3.1 Correlation Between Knowledge of Nanotechnology, Religiosity, and Mean Ethical Concern

To examine the relationship between the following variables, a Pearson product-moment correlation was conducted. This regression test measures the strength of a linear relationship between two variables and uses a coefficient from -1 to 1, with -1 being a strong negative linear correlation and +1 being a strong positive correlation. Among all the variables, there was a statistically significant relationship between knowledge of nanotechnology and mean ethical concern ($p < 0.05$) and a negative association between the two ($r = -0.185$). The findings suggest that a 1-point increase in knowledge of nanotechnology meant a 0.185-point decrease in overall ethical concern, or that more knowledge led to less concern for the usage of nanotechnology. Since the r-value was miniscule, there was only a slight correlation between the two variables.

The relationship between knowledge of nanotechnology and religiosity scores had a p-value = 0.26 and Pearson’s r-value = -0.089 indicating that even though the data was not as significant, there was a negative correlation between the two meaning that a 1-point increase in the knowledge of nanotechnology had been linked to a 0.089-point decrease in religiosity. On the other hand, religiosity and the mean ethical concern were positively correlated with a p-value = 0.050 and Pearson’s r-value = 0.157, showing that a 1-point increase in religiosity was linked to a 0.157-point increase in ethical concern. Both scatter plots had small slopes because the range of numbers used to represent each variable did not go past 5. However, there was not an extreme relationship between the two. A list of the correlations is shown in Table 11 and scatterplots of all variables are shown in Figure 5.

Table 11. Pearson’s correlation of all the variables

	n	Pearson’s r	p	VS-MPR †
Knowledge of Nanotech - Mean Religiosity Scores	157	-0.089	0.266	1.045
Knowledge of Nanotech – Mean Ethical Concern	157	-0.185*	0.020	4.690
Mean Religiosity Scores – Mean Ethical Concern	157	0.157	0.050	2.455

*p < 0.05

Scatter plots

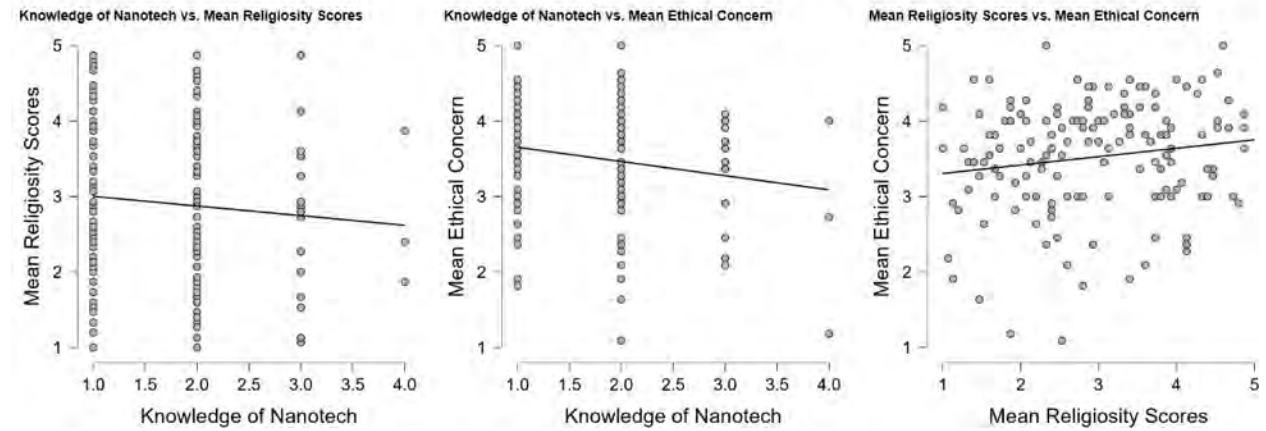


Figure 5. Scatter plots of knowledge of nanotechnology, religiosity, and ethical concern correlated with each other.

Finally, age was also used as a variable to see if any of the prior variables had shown any relationship with age. As shown in Table 12, age was most strongly correlated with knowledge of nanotechnology, depicting that a 1-point increase in age was linked to a 0.201 increase in knowledge of

nanotechnology. The p-value = 0.011 which is very close to 0.01 meaning that the data is relatively significant. The Pearson's r-value = 0.201 which represents a slight positive correlation between the two variables.

Table 12. Pearson’s correlation of all the variables including age

	Pearson’s r	p	Lower 95% CI	Upper 95% CI
Knowledge of Nanotech - Mean Religiosity Scores	-0.074	0.351	-0.226	0.082
Knowledge of Nanotech – Mean Ethical Concern	-0.176*	0.026	-0.322	-0.022
Mean Religiosity Scores – Mean Ethical Concern	0.137	0.083	-0.018	0.286
Ages – Knowledge of Nanotech	0.201*	0.011	0.048	0.345
Ages – Mean Ethical Concern	0.033	0.682	-0.123	0.186
Ages – Mean Religiosity Score	0.012	0.879	-0.143	0.166

* p < 0.05

4. Discussion

Upon analyzing the statistical results found, there was only a statistically significant difference between religiosity and the religions mentioned, a slight negative correlation between knowledge of

nanotechnology and mean ethical scores, and a positive correlation between increasing age and knowledge of nanotechnology. Primarily, the findings suggest that the religions tested have differing religiosity scores which may play a role in their connection with science and moral understanding of

technology (Franz Foltz and Fredrich Foltz, 2006). Additionally, Brossard, Dominique et al. 2009, had studied that lower knowledge of nanotechnology meant lower reasons to support it, which agrees with the data as there was a negative correlation between the two. However, unlike the results of this study, her study utilized a hierarchical ordinary least squares (OLS) regression to find that lower knowledge of nanotechnology was often paired with higher levels of religiosity. In this study, there was a small negative correlation between knowledge and religiosity which was consistent, but the correlation was too close to zero to say it was significant. The observations found from the other variables including religiosity vs. ethical concern (positive correlation) also agreed with previous studies, but the correlation coefficient was not large enough to show a prominent relationship.

Overall, through ANOVA, this study identified that there were statistically different religiosity scores present in each religion, but there was no difference in knowledge scores across those religions. By doing a Pearson's Correlation, there only seemed to be slight evidence pointing to an agreement in previous trends among my variables, with the strongest being knowledge vs. mean ethical concern. The ambiguous and weak relationships among students compared to adults may suggest that adolescents use more affective or emotional factors rather than cognitive and sociocultural factors when identifying their perceived ethical concern toward nanotechnology, according to Nano- PAAF (Pillai & Bezbaruah, 2017). A lack of understanding about nanotechnology, despite including an information page, may have also contributed to their limited perception of nanotechnology, hence, accounting for the negative correlation between knowledge and overall ethical concern.

When looking further into the potential ethical dilemmas, nanotechnology getting into the "wrong hands" and the loss of freedoms and privacy seemed to be the biggest issue concerning adolescents rather than the social footprint or even the long-term health effects of nanotechnology. This suggests that teens care more about drastic future implications or a sense of "evil" that may polarize the use of nanotechnology, instead of direct implications. Contrasting the findings, Macoubrie (2006) had

found that the adults in her study were concerned about the long-term health issues and military uses of nanotechnology, while even the social footprint category appeared to have a high frequency in her study. Each ethical concern was also correlated and there was mostly a significant positive association among all of them, with loss of freedoms vs. lack of regulation having the highest p-value and thus, having the strongest correlation.

4.1 Implications

Although some results found in this study were not significant enough to make a conclusion, the general trends do show some similarities between prior studies, suggesting a need to further explore the relationships between adults and adolescents within their understanding of nanotechnology and the ethical concerns it poses. Since there was a clear difference in ethical concerns that students prioritized, future research can help address those issues in ways that specifically cater to their knowledge of nanotechnology and religiosity. The nanotechnology engagement project can also be expanded into classroom settings since the lack of knowledge found by students in Granada Hills poses the need to introduce modern technological advancements like nanotechnology into the school curriculum (Pidgeon, Nick et al. 2011). The results of this study provide critical information about the relationship between religion, religiosity, prior knowledge, and ethical concerns of nanotechnology that may be used to drive future studies on greater populations outside Granada Hills, California.

4.2 Limitations

While conducting research, I tried obtaining a representative sample by contacting an equal number of teachers from both schools. However, only minimal responses came from School 2, which led to underrepresentation in data from School 2. Although both schools have similar demographics, since they are located in the same area, the unequal distribution of respondents may have skewed the results and led to a less reliable generalization. Additionally, some religions were not included in the statistical analysis

(Mormons, Jews, etc.) because of the lack of data supported by those groups. Therefore, the data may not be completely representative of communities with higher demographics of those religions, and further research must be done to thoroughly analyze those communities.

Another limitation is that the data collection instrument could have been modified better. Primarily, in the Google Form, the information page that discussed the practical and societal uses of nanotechnology was followed right after the section that inquired about the respondent's prior knowledge of nanotechnology. Therefore, participants could have easily changed their responses after knowing the premises of nanotechnology, skewing the data about their knowledge of nanotechnology. Secondly, the nanotechnology information page could have been adjusted to include links and resources that the students could go to if they did not understand what was on the page, as many participants in the pilot test overlooked or could not understand the nanotechnology information page properly. With a well-written information page, students could have had a clear understanding of the uses of nanotechnology and properly identified what ethical concerns meant the most to them instead of randomly inferring.

Finally, Nano-PAAF stresses the importance of other factors that go into the perceptions of nanotechnology, other than the ones studied in this paper (Pillai & Bezbaruah, 2017). These factors could have an equal or more influential impact on students' ethical concerns about nanotechnology, such as their political views, gender, intake of scientific media, affective values, etc. This study did not analyze these factors because some are more subjective and harder to quantify, while others cross the ethical boundaries of the participants.

5. Conclusion

Overall, this study examined a few of the several factors that affect people's perceptions of nanotechnology and found a relationship between religiosity and religion, the various ethical concerns, and weaker correlations between knowledge,

religiosity, and ethical concern. The gap was further analyzed as I concluded that adolescents did have similar trends between the variables as adults, except their ethical concerns were centered more around a sense of "evil" or drastic polarization of nanotechnology use. Therefore, the study's findings can be used to create more educational programs around the acceptance and integration of nanotechnology, so that students are less fearful and unaware of nanotechnology. The worldwide acceptance of technology can be improved by examining religion and religiosity further and adapting products to address the ethical concerns of people from different religions.

In the future, more research should be conducted on adolescents in different areas of the United States using a greater sample population since the rising technology industry will be directed and established by the newer generations. To build on this research, the affective factors (ex. hope, expectations, fears, feelings) should be correlated with the variables mentioned in this paper to improve the knowledge known about students and their growing perceptions of nanotechnology. Finally, a focus group could be beneficial for conducting more research since students sometimes feel more comfortable and informed around their own peers who are all on the same page, rather than individually.

References

- A., Wang S. (eds) *Nanotechnology in Dermatology*. Springer, New York, NY. https://doi-org.oca.ucsc.edu/10.1007/978-1-4614-5034-4_25
- Am T. G. (2011). Trust in Nanotechnology? On Trust as Analytical Tool in Social Research on Emerging Technologies. *Nanoethics*, 5(1), 15–28. <https://doi.org/10.1007/s11569-010-0105-8>
- Binder, A.R. (2013). Understanding Public Opinion of Nanotechnology. In: Nasir, A., Friedman, A., Wang, S. (eds) *Nanotechnology in Dermatology*. Springer, New York, NY. https://doi.org/10.1007/978-1-4614-5034-4_25
- Brossard, Dominique, Dietram A Scheufele, Eunkyung Kim, and Bruce V Lewenstein. (2009).

Religiosity as a Perceptual Filter: Examining Processes of Opinion Formation About Nanotechnology. *Public understanding of science (Bristol, England)* 18, no. 5, 546–558.

Diamant, J., & Sciupac, E. P. (2020, September 10). *10 key findings about the religious lives of U.S. teens and their parents*. Pew Research Center, <https://www.pewresearch.org/fact-tank/2020/09/10/10-key-findings-about-the-religious-lives-of-u-s-teens-and-their-parents/>

Florczyk SJ, Saha S. (2007) Ethical issues in nanotechnology. *J Long Term Eff Med Implants*. 17(3):271-80. doi: 10.1615/jlongtermeffmedimplants.v17.i3.90. PMID: 19023950.

Foltz, Franz A, and Frederick A Foltz. (2006) The Societal and Ethical Implications of Nanotechnology: A Christian Response. *The Journal of technology studies* 32, no. 1/2 104–114.

Ganesh Pillai, R., & Bezbaruah, A. N. (2017). Perceptions and attitude effects on nanotechnology acceptance: An exploratory framework. *Journal of Nanoparticle Research*, 19(2). <https://doi.org/10.1007/s11051-016-3733-2>

Gardner, Grant, Gail Jones, Amy Taylor, Jennifer Forrester, and Laura Robertson (2010) Students' Risk Perceptions of Nanotechnology Applications: Implications for Science Education. *International journal of science education* 32, no. 14, 1951–1969.

Gavas, S., Quazi, S., & Karpiński, T. M. (2021). Nanoparticles for Cancer Therapy: Current Progress and Challenges. *Nanoscale research letters*, 16(1), 173. <https://doi.org/10.1186/s11671-021-03628-6>

Good, M., Willoughby, T., & Busseri, M. A. (2011). Stability and Change in Adolescent Spirituality/Religiosity: A Person-Centered Approach. *Developmental Psychology*, 47(2), 538–550. <https://doi.org/10.1037/a0021270>

Huber, Stefan, and Odilo W. Huber 2012. "The Centrality of Religiosity Scale (CRS)" *Religions* 3, no. 3: 710-724. <https://doi.org/10.3390/rel3030710>

Institute of Medicine (US) Food Forum. Nanotechnology in Food Products: Workshop

Summary. Washington (DC): National Academies Press (US); 2009. 4, Educating and Informing Consumers About Applications of Nanotechnology to Food Products. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK32728/>

Isabella A. Joubert, Mark Geppert, Stefanie Ess, Reinhard Nestelbacher, Gabriele Gadermaier, Albert Duschl, Arne C. Bathke, Martin Himly, Public perception and knowledge on nanotechnology: A study based on a citizen science approach, *NanoImpact*, Volume 17, 2020, 100201, ISSN 2452-0748, <https://doi.org/10.1016/j.impact.2019.100201>.

Kotze, Manitza. (2018) The Theological Ethics of Human Enhancement : Genetic Engineering, Robotics and Nanotechnology. *In die skriflig : tydskrif van die Gereformeerde Teologiese Vereniging* 52, no. 3, 1–8.

Loskota, B. (2012, November 9). *Appendix: Geographic profile of Los Angeles*. Appendix: Geographic Profile of Los Angeles | Center for Religion and Civic Culture. <https://crcc.usc.edu/report/faithful-action-working-with-religious-groups-in-disaster-planning-response-and-recovery/appendix-geographic-profile-of-los-angeles/>

Macoubrie, Jane. (2006) Nanotechnology: Public Concerns, Reasoning and Trust in Government. *Public Understanding of Science* 15, no. 2, 221–41. <https://doi.org/10.1177/0963662506056993>.

Nasrollahzadeh, Mahmoud & Sajadi, S.Mohammad & Sajjadi, Mohaddeseh & Issaabadi, Zahra. (2019). Applications of Nanotechnology in Daily Life. 10.1016/B978-0-12-813586-0.00004-3.

Patten, M. L., & Newhart, M. (2018). *Understanding research methods: An overview of the Essentials*. Routledge.

Patil, M., Mehta, D. S., & Guvva, S. (2008). Future impact of nanotechnology on medicine and dentistry. *Journal of Indian Society of Periodontology*, 12(2), 34–40. <https://doi.org/10.4103/0972-124X.44088>

Peter D. Hart Research, Inc. (2008) Awareness of and Attitudes Toward Nano-technology and Federal Regulatory Agencies. Available online at

http://www.pewtrusts.org/our_work_report_detail.aspx?id=30539.

Pidgeon, Nick F., Barbara Herr Harthorn and Terre Satterfield (2011) Nanotechnology risk perceptions and communication: emerging technologies, emerging challenges. *Risk analysis : an official publication of the Society for Risk Analysis* 31, 1694-700 .

Scheufele, Dietram A., et al. (2009) Religious Beliefs and Public Attitudes toward Nanotechnology in

Europe and the United States. *Nature Nanotechnology*, vol. 4, no. 2, pp. 91–94. *EBSCOhost*, doi:10.1038/nnano.2008.361.

Toumey, C. (2004). Narratives for Nanotech: Anticipating Public Reactions to Nanotechnology. *Techné: Research in Philosophy and Technology*, 8, 88-116.

What is nanotechnology? National Nanotechnology Initiative. (n.d.), <https://www.nano.gov/nanotech-101/what/definition>

Comprehensive Review on Atopic Dermatitis and Therapeutic Potentials

Crystal Shin^{1*}

¹Northwood High School, Irvine, CA USA

Received September 3, 2022; Revised March 7, 2023; Accepted, March 20, 2023

Abstract

Atopic dermatitis (AD) is a type of a skin-related allergy that affects all demographics worldwide, but mostly common amongst the younger population. It is characterized by itchy, red, and swollen skin, and there is currently no cure available. Disease etiology largely arises due to a combination of genetics and environmental influences, coupled with other types of comorbidities such as asthma and hay fever disease. AD largely is associated with a compromised epithelial tissue layer, which leads to decreased water retention and facilitated entry of allergens into the body, which further exacerbates the pathophysiology of the disease. Mechanistically, the progression of AD can be attributed to the overall imbalance of Th1 and Th2 cells, which disturbs the homeostasis between pro-inflammatory and anti-inflammatory polarizations of innate and adaptive immunity. Therefore, most of the current medications that help ameliorate AD's pathogenesis include drugs that inhibit Th2 differentiation and polarization by targeting the activity of IL-13 and IL-4, both of which are implicated in the production of Th2 cells that subsequently induce IgE-mediated allergic reactions. The purpose of this review article is to comprehensively explore the biology of AD, including risk factors, diagnosis, pathophysiology, and the current available treatments.

Keywords: Atopic dermatitis, Allergy, IgE, Inflammation, Hypersensitive immune disorder, Treatment

1. Introduction

Atopic dermatitis (AD) is a chronic form of skin-related inflammation that results in itchy, red, swollen, and cracked skin (U.S. Department of Health and Human Services, 2022). It currently affects approximately 2.4% of the global population, impacting around 15-20% of children and 1-3% of adults worldwide. AD is neither a fatal nor contagious condition, but its Incidence has grown over 3-folds since the 1970s, and the prevalence is expected to rise over the next decade. It affects nearly all demographics, although recent evidence indicates that AD is commonly present in younger individuals

under the age of 5 (Woods, 2017).

AD is classified under a broad category of an inflammatory skin condition known as eczema. Eczema can be broadly divided into two types: Atopic dermatitis (which is the focus of this paper) and contact dermatitis (WebMD, 2022). In brief, contact dermatitis, as the name implies, results when a skin comes into contact with an external trigger that causes a rash. Depending on the trigger, contact dermatitis can be grouped into two types: Irritant dermatitis and allergic dermatitis. Irritant dermatitis is much more common, and its triggers include cosmetic products, soaps & detergents, nickel-based jewelry, and industrial chemicals like cement

* Corresponding Author
crystalart0305@gmail.com

Advisor: Min Sub Lee
minsub.lee@ucla.edu

(WebMD, 2022). Allergic dermatitis on the other hand, flares when a skin comes into contact with an allergen that an individual is allergic to. Common allergens include pollen, poison ivy, rubber, latex, fragrances, and in extreme cases, water and sunlight.

Atopic dermatitis on the other hand, is a chronic condition often due to a combination of hereditary, immune, and environmental factors (Tee-Melegrito, 2021). AD can also occur in individuals with a family history of asthma or hay fever disease. The purpose of this review article is to provide a comprehensive background on AD, starting with the risk factors, major causes of the disease, and current diagnosis. It will then discuss the pathophysiology of the disease, including some background in immunology as well as the molecular mechanisms governing disease progression. Finally, the paper will conclude with an insight on some of the currently available therapeutics and treatments for AD.

2. Risk Factors

As with many types of inflammatory pathologies, there is no direct root cause of AD. Instead, it results from a culmination of multiple factors, such as genetics, environmental, and immunological factors. This section breaks down some of the major risk factors that contribute to the development of AD in most individuals.

2.1 Family History of Atopy

Atopy is a type of hypersensitive immune disorder characterized by the production of exaggerated amounts of immunoglobulin E (IgE) to harmless substances in the environment (i.e., non-microbial food or airborne allergens, such as dust mites, pollen, and foods) (Watson, 2021). Upwards of 70% of patients with AD also have a family history of atopy (Woods, 2017). In addition, some studies point out that individuals carrying AD have contracted other atopic diseases prior to acquiring AD, such as food allergies, asthma, and allergic rhinitis, in sequential order (Hahn and Bacharier, 2005). Although these atopic conditions are not necessary precursors for developing AD, it isn't uncommon for an individual to carry some

combination of these pathologies (e.g., A patient with AD is also associated with having asthma, food allergy, allergic rhinitis, or a combination of all three).

2.1 Genetic Mutations

Growing evidence increasingly suggests that genetic factors play a role in development of AD, specifically via a loss-of-function mutation in a gene known as *flg*. *Flg* encodes for a protein called filaggrin, whose function is to regulate epidermal homeostasis by bringing together keratin fibers and other structural proteins together to establish a strong epithelial barrier on the skin surface (Sandilands, et al., 2009). Filaggrin deficiency leads to a “leaky” epithelial skin barrier that leads to decreased water retention and facilitated entry of allergens through the epidermis (McLean, 2011). Because of the increased water loss, individuals containing *flg* mutation tend to have dry, scaly skin, and they are much more susceptible to contracting AD due to increased exposures to environmental allergens (McLean, 2011). Defects in *flg* have also been associated with asthma, which further validates the propensity of individuals simultaneously carrying both asthma and AD (Woods, 2017).

While *flg* plays a pivotal role in maintaining the integrity of the epithelial barrier, it is certainly not the only involved protein. In fact, only about 30% of European patients with AD carry a *flg* mutation, suggesting that other epithelial gene variants are yet to be identified (Woods, 2017). Other skin barrier factors include a lack of ceramide production (Yang, et al., 2020). Ceramides are essential lipid molecules that are localized in the plasma membrane of epithelial cells that help regulate the biochemical & biophysical properties of the phospholipid bilayer. A reduction in ceramide concentration in the plasma membrane is associated with a loss of tight junctions that are critical for maintaining the epithelial network (Jennemann, et al., 2007). While no research has specifically investigated the direct relationship between *flg* mutation and a loss of ceramide production, it would be unsurprising to find that *flg*

expression and ceramide biosynthesis could operate in a shared biological pathway.

2.3 Hygiene Hypothesis

Why is it important for kids to play outside while they're young? According to the hygiene hypothesis, childhood exposure to various types of microorganisms is vital for protection against future allergic responses via immune system development (FDA, 2018). The biological basis of the hygiene hypothesis is a disequilibrium between Th1 and Th2 helper T-lymphocytes. Th1 immunity is associated with proinflammatory responses that primarily defends the body against intracellular pathogens (e.g., viruses and parasites) and stimulates autoimmune responses via production of a cytokine, interferon gamma (Berger, 2000). On the other hand, Th2 immunity stimulates the humoral immune response by stimulating B-cell proliferation, differentiation, and antibody production via a cytokine, interleukin-4 (IL-4).

Allergic conditions are caused by inappropriate immunological responses to harmless antigens driven by Th2-mediated immunity. This is because Th2 cells produce IL-4, IL-5, and IL-13, which predominantly stimulates B-cells to produce large quantities of IgE that subsequently activate mast cells (Okada et al., 2010). In alignment with this biological phenomenon, it has been shown that skin lesions of patients carrying AD contains high infiltration of Th2 population of cells relative to other T-helper subtypes, suggesting a link between Th2-dependent stimulation of humoral immunity and IgE-mediated allergic response (Brandt and Sivaprasad, 2011).

Furthermore, the hygiene hypothesis is vital for the development of regulatory T-cells (Tregs). The function of Tregs is to suppress overactive immune response, thereby maintaining immune homeostasis and self-tolerance (Kondelkova, et al., 2010). Individuals with inadequate Tregs response are more susceptible to autoimmune diseases and hypersensitive disorders (e.g., allergies) because of insufficiently repressed Th1 and Th2 cells, respectively (Bufford and Gern, 2005).

Altogether, early childhood exposures to

microbes can substantially reduce the likelihood of individuals contracting AD and other forms of allergies. Some studies suggest that children who are exposed to dogs are correlated with lower risk of developing AD, likely because dogs can significantly increase a young child's exposure to environmental pathogens (Pelucchi, et al., 2013). Likewise, children with poor hygiene and those who consume unpasteurized milk tend to have a lower risk of developing AD (Flohr and Mann, 2014). Of course, this does *not* imply that all children should abandon practicing personal hygiene and sanitation.

2.4 Nutritional and Dietary Factors

Another fundamental root cause of AD can be attributed by dietary influences, starting as early as fetal development during pregnancy and breastfeeding. This is because diet & nutrition is crucial for the development of a healthy gut microbiome, which can subsequently influence the integrity of the skin barrier. One study has shown that children born to mothers with low vitamin D intake during pregnancy have an increased risk prevalence of AD (Miyake, et al., 2010). Interestingly, it has also shown that children born during fall and winter tend to have greater association with AD compared to children born during spring and summer. This might be attributed to the fact that there is less sunlight exposure during fall and winter, which correlates to lower exposure of vitamin D (Kuzume and Kusu, 2007).

While the relationship between breastfeeding and allergic risk hasn't fully been determined, breastfeeding has been associated with lower incidences of allergic diseases, including AD (Minniti, et al., 2014). This is because breastfeeding supports diversification of microbial colonization as the infant acquires some of their mother's healthy microbes. A healthy gut microbiome is vital for maintaining a well-balanced immune system. Therefore, breastmilk can potentially reduce risk of developing AD via increased diversification of an infant's microbiome compared to intaking baby formulas, although more follow-up studies are needed to validate these early findings (Trikamjee, et al., 2021).

In addition, several studies have shown that early dietary supplementation of long-chain polyunsaturated fatty acids (LCPUFA) in the form of fish oil has shown reduced incidence of allergic diseases. Mechanistically, high amounts of LCPUFA in the plasma membrane of cells reduce the synthesis of prostaglandin E and subsequent inhibition of cytokine and IgE production (Trikamjee, et al., 2021). As stated previously, reduced IgE production is a proxy for reduced allergic diseases, including AD.

2.5 Climate and Environment

Climate and environment can also play a vital role in the development of atopic dermatitis. Winter seasons are characterized with low humidity and low temperatures, and thus, it is common for individuals to have cracked skins due to loss of moisture on the skin surface. Perhaps unsurprisingly, this increases the risk of individuals contracting AD because of the compromised skin barrier functions and increases susceptibility to mechanical stress (Engebretsen, et al., 2016). In addition, there are many environmental toxins and carcinogens that are associated with the acceleration of eczema and atopic dermatitis. These include: Outdoor pollutants (e.g., carbon monoxide, dusts, forest fires, industrial wastes from power plants, etc.), tobacco smoke, and indoor carcinogens (e.g., paint, high exposure to cosmetic products, burning stoves, cleaning products, etc.) (Shi, 2022). All of these products trigger inflammatory reactions on the skin, thus contributing to the loss of skin barrier function and rendering individuals susceptible to increased water loss and subsequent xerosis.

2.6 Final Thoughts

As with many disease models, it's difficult to ascertain whether these risk factors independently spearhead AD development or depend on each other to collectively cause AD. This is largely because of individual differences in genetics, gender, ethnicity, lifestyle, diet, and many more. In some cases, a genetic mutation alone (or a combination of

multiple mutations) is sufficient to drive AD development. In other cases, a combination of multiple environmental factors (i.e., climate, diet, pollutants, etc.) may be necessary to exacerbate the pathophysiology of AD but not sufficient to initiate AD development. Therefore, more studies are needed to better evaluate the underlying root causes of AD and the biological basis of its persistence.

3. Diagnosis

The Hanifin and Rajka criteria are the most notably recognized set of diagnostic criteria used in modern-day medicine to screen patients for AD. It is clinically diagnosed based on symptoms alone, without necessarily performing rigorous testing (Woods, 2017). As of 1980, Hanifin and Rajka standards have categorized AD diagnosis requirements into two sections: Major and Minor. The major features of Hanifin and Rajka in determining the presence of AD include: (1) Pruritis and eczema, both of which are characterized by itchy and dry skin, rashes, scaly patches, and blisters; (2) Typical morphology and distribution; (3) Chronically relapsing dermatitis; and (4) Personal or family history of atopy (asthma, allergic rhinitis, and atopic dermatitis). Currently, there are 23 minor features, such as elevated immunoglobulin E (IgE) levels, food tolerance, tendency towards cutaneous infections, anterior neck folds, and more. These guidelines have proposed to diagnose a patient with AD when they meet at least three of the four major features *and* at least three of the minor features (Akan, et al., 2020).

One ongoing challenge of AD diagnosis is the fact that its symptoms overlap with other related skin diseases. Not only are there several imitators with similar clinical features, but there are many diagnoses that coexist with AD. For example, seborrheic dermatitis is another type of inflammatory skin disease that is very difficult to distinguish from AD, especially in infancy where they can occur concomitantly or separately (Woods, 2017). However, there *are* a few distinguishing features specific to AD. In nearly most cases, AD begins during infantile onset, while other forms of skin

diseases such as contact dermatitis are rarely found in infants. Furthermore, AD does not consist of circumscribed lesions but instead, consist largely of xerosis. This is in contrast with both seborrheic and contact dermatitis, as both skin diseases consist of circumscribed lesions (Siegfried, 2015). Finally, another unique aspect of AD is its distribution in the body at various developmental stages. For infants, AD is largely found in the face and the trunk. In children however, AD is localized near the flexors (e.g., hip, thigh, and torso), while in adults, AD is found mostly in the hands (Siegfried, 2015). Nevertheless, it isn't uncommon for patients to be diagnosed with AD in conjunction to other skin diseases due to large overlaps in symptoms, distributions, and diagnosis.

4. Pathophysiology

At the end of the day, atopic dermatitis is only one specific type of a skin allergy, amongst numerous other pathologies. Most allergic reactions are governed by the secretion of IgE antibodies from mature B-cells, which in turn, activates mast cells to secrete high amounts of histamine and granules that contribute to hyperinflammatory reactions in specific types of organs. This section discusses broad cellular and molecular mechanisms that lead to an allergic reaction and then focus on the specific signaling pathways that become activated during AD. Understanding these processes is crucial because it enables researchers to design appropriate drugs that target and combat disease progression.

4.1 Innate vs. Adaptive Immunity, Antigen Presentation, and IgE Antibody

As described in the above sections, all allergic reactions begin when a non-microbial allergen enters the body and the immune cells *fail* to distinguish an allergen from a real microbe. When a real microbe enters the host tissues, the innate immune cells immediately recognize the foreign microbe as a threat and assemble together to destroy the pathogen. All microbes contain specific elements known as pathogen-associated molecular patterns (PAMPs), and different types of PAMPs

are recognized by different innate immune cells via pathogen recognition receptors (PRRs) (Mogensen, 2009). For example, some gram-negative bacterial cells contain a specific toxin within its cell wall known as lipopolysaccharide (LPS). Over evolutionary time, eukaryotic host cells have evolved a receptor known as Toll-Like Receptor (TLR)-4 that specifically recognizes secreted LPS molecules from gram-negative bacteria and subsequently activates inflammatory reactions that enable the host organism to fight against the invading bacteria (Raetz and Whitfield, 2002). In order to maintain long-term memory response against that bacteria, the innate immune system must be able to transmit these signals to the adaptive immune system (consisting of T and B-cells) that generates memory lymphocytes.

Dendritic cells (DCs) are the major type of immune cell that bridges the innate and adaptive immune system. When DCs encounter a foreign pathogen, it usually ingests that microbe via phagocytosis ("cellular eating"), processes that microbe into tiny peptides, and presents those peptides as an antigen in the context of a molecule known as major histocompatibility complex (MHC). Once the DCs present these foreign antigens, circulating naive T-cells (immature T-cells that haven't encountered any antigen) will come into contact with the MHC molecules containing the antigen and subsequently become activated into a mature T-cell (They and Amigorena, 2001). These mature T-cells are able to secrete cytokines and other signaling molecules that lead to the proliferation and expansion of an antigen-specific T-cell population that can fight off any remaining pathogens in the body. Once the pathogen is sufficiently cleared, most of these T-cells undergo apoptosis, while some T-cells differentiate into memory T-cells (Harrison, et al., 2019). Memory lymphocytes will re-differentiate into effector T-cells when the same pathogen enters the body, resulting in facilitated pathogen clearance from the host.

The activated T-cells secrete cytokines that also lead to maturation and differentiation of native B-cells into plasma cells, which is the immune system's factory machine that secretes huge amounts of antibodies (Harrison, et al., 2019).

There are 5 main types of antibodies (IgA, IgD, IgE, IgG, and IgM), with each eliciting different effector functions, although the function of IgD isn't well-characterized (Hoffman, et al., 2016). As described in previous sections, IgE is the primary antibody that is responsible for regulating allergic reactions.

4.2 Molecular Mechanisms of Allergic Reactions

An allergic reaction begins when an allergen enters a body and triggers the DCs to uptake the allergen and present it via MHC molecules to circulating naive T-cells, either naive CD8 T-cells or naive CD4 T-cells. Activated CD8 T-cells move on and become engaged in cytotoxic effector functions (e.g., target and kill tumor cells) while activated CD4 T-cells can differentiate into specific subtypes depending on the cytokine exposures: Th1, Th2, Th17, and Tregs (Harrison, et al., 2019). Th2 subtype is most directly engaged in allergic reactions, and thus, it will be the main focus for this section. Th2 cells produce a cytokine called interleukin 4 (IL-4), which directly stimulates B-cells to produce large amounts of IgE antibodies. These IgE antibodies begin to circulate throughout the bloodstream and bind to IgE-receptors that are typically found on the surface of mast cells and basophils (Stone, et al., 2010). It's also important to note that these IgE antibodies are highly specific to the allergen that was originally presented by the DCs into naive T-cells.

When the individual encounters the same allergen the second and subsequent times, the allergen binds to the IgE antibodies that are coated on the surface of the aforementioned mast cells and basophils. This binding reaction activates the mast cells and basophils to undergo a process known as degranulation, which involves a massive release of histamines and other inflammatory mediators from their stored granules (Stone, et al., 2010). These inflammatory processes lead to vasodilation, rise in body temperature, localized and systemic swelling, and mucous secretion, all of which are common physiological symptoms during acute infection. Depending on the type of allergen, the individual's immune response, and the mode of allergen

consumption, allergic symptoms can vary from localized inflammatory reactions to systemic anaphylactic shock (Amin, 2012).

4.3 Pathogenic Insights of Atopic Dermatitis

Systemic secretion of histamines and other inflammatory mediators via mast cells and basophils represents an "umbrella" of all allergic reactions. Many studies have characterized other molecular, cellular, and physiological defects that contribute to the exacerbation of AD-mediated skin inflammation. The previous section has already identified *flg* mutation as a risk factor of AD because the individual has a genetic defect in skin barrier function.

Recent studies have found a novel link between NOD-like receptors (NLRs) and susceptibility to atopic dermatitis. NLRs are intracellular PRRs that sense a wide range of PAMPs such as bacterial cell wall's peptidoglycans and various derivatives of amino acids that are not resident in host organisms (Tsang, et al., 2021; Zhong, et al., 2013). Nlrp12-null genetic models, which are deficient in NLR sensing, have been shown to exhibit attenuated atopic dermatitis-like features, as well as reduced antigen presentation and subsequent T-cell activation (Arthur, et al., 2010). This suggests that the activation and progression of AD could possibly operate through NLR signaling. NLR signaling also leads to the activation of the inflammasome in keratinocytes as well as many residential innate immune cells in the skin, which leads to the production of the cytokines, IL-1 β and IL-18 (Jacobs and Damania, 2012). These cytokines in turn can trigger Th2 and IgE-mediated immune responses, which contribute to the activation of mast cells and basophils.

In parallel to these studies, it is also known that patients who have a preliminary exposure to *Staphylococcus aureus* or herpes simplex virus exhibit increased vulnerability to AD. Products of these pathogens have been shown to downregulate immune responses by downregulating the FcR1 receptor on Langerhan cells upon binding to TLR-2 receptor, resulting in poor TH1 activation but enhanced TH2 response (Schlapbach and Simon,

2014).

Another cytokine that has been shown to trigger pathogenesis of AD is thymic stromal lymphopoietin (TLSP), which is secreted by epithelial cells. The epidermis of the lesions derived from AD patients has been shown to express greater levels of TLSP compared to control patients with nonallergic dermatitis (Soumelis, et al., 2002). Dendritic cells from these lesions show an activated phenotype, in which they migrate away from the epidermis and into the lymph nodes to directly promote Th2 differentiation. This is consistent with the observation that a direct exposure of TLSP enables DCs to activate naive CD4 T-cells and differentiate them into Th2 subtype (Indra, 2013). Furthermore, it is thought that the high levels of cytokine production (e.g., IL-1 β , TNF α , IL-4 and IL-13) can synergize to stimulate mass production of TLSP from surrounding epithelium and keratinocytes, suggesting a positive feedback loop of inflammatory circuit within the skin (Indra, 2013). However, these studies are correlation at best because no loss-of-function studies have been performed in the context of TLSP signaling, and thus, more studies are needed to validate this inflammatory circuit between TLSP and DC-dependent Th2 differentiation.

Lastly, Interleukin (IL)-33 has been shown to be overexpressed in keratinocytes derived from patients with AD, suggesting that IL-33 may be involved in triggering AD development (Imai, 2019). IL-33 stimulates a wide range of immune cells, including group 2 innate lymphoid cells (ILC-2) and basophils, both of which produce IL-4, which triggers a Th2 response and subsequent production of IgE antibodies that lead to progression of AD (Imai, 2019). It is also important to note that IL-33 reduces the expression of filaggrin, which as described previously, is integral for maintaining the stability of the epithelial barrier (Imai, 2019). The circuit between IL-33 and filaggrin hasn't fully been explored in AD patients, and thus, it would be an interesting therapeutic target once the cellular and molecular mechanisms are fully deciphered.

5. Current Treatments

This last section will explore different strategies of combating and treating atopic dermatitis. As for many diseases, there is currently no available cure of AD. But with the right treatments, the severity of flares and inflammation can be reduced quite substantially.

5.1 Lifestyle Changes

Getting rid of a disease does not happen overnight. If there were such treatments, everyone on Earth would be disease-free by now. Part of the reason why many patients are unable to recover from certain types of disease is because they lack self-perseverance and consistency. For example, if an individual is diagnosed with obesity, one of the most definite ways of recovering from obesity is to amend diet and exercise every day. Of course, this is grossly oversimplifying the concept because not all obesity conditions can be combated via lifestyle changes. However, many studies have proved over and over that a 30-minute aerobic exercise coupled with a high protein diet for a month is sufficient to shed anywhere between 10-15 pounds in a month (Goldman, 2017).

Likewise, AD can be treated as long as patients maintain consistent routines and adapt to new changes whenever necessary. One daily routine that can help alleviate AD is to bathe in lukewarm and salty water. This can help moisturize skin by restoring the protective barrier of the skin and maintaining proper levels of hydration (Lio, 2013). Saltwater also contains minerals such as potassium, calcium, and magnesium, all of which can bind to water and help with retention of moisture in the skin (NENA skincare, 2022).

Another interesting lifestyle method would be to avoid woolen clothing and instead, wear more silk-based clothing (NENA skincare, 2022). It has been shown that the sericin residue found in silk has a repellent property that prevents the attachment of dust mites, bacteria, and other common allergens (Silk Properties, 2022). Health professionals often advise AD patients to avoid woolen clothing because wool tends to remove moisture from the body, which

can exacerbate the inflammatory reactions on the skin surface (NENA skincare, 2022).

Finally, while this one isn't necessarily a "lifestyle change," it can be a very useful device to help reduce the severity of AD: Humidifier. Winter season can be very dry due to the lack of humidity in the air. A humidifier is an electrical appliance that can help increase humidity in the room, which can significantly prevent skin from drying out and consequently reduce risk of contracting AD (Varothai, et al., 2013). On top of that, consistently applying moisturizers and practicing proper skin hydration routines can also help prevent the progression of AD.

5.2 Medication

Topical treatment of hydrocortisone cream is a widely popular medication used for reducing AD. Hydrocortisone is a corticosteroid that is used as an agonist for glucocorticoid and mineralocorticoid receptors that help suppress inflammatory reactions in the tissues via inhibition of inflammatory mediators, such as phospholipase A2 and NF-kappa-B (National Center for Biotechnology Information, 2022). It was first patented and approved for medical use by 1941, and as of 2019, it was listed as the 147th most commonly prescribed medication in the United States (Kane, 2022).

Tacrolimus (aka Prograf) is a type of ointment that can also be topically treated onto allergic sites (Mayo Foundation for Medical Education and Research, 2022). Mechanistically, it acts as a calcineurin inhibitor in T-cells. When a naive T-cell comes in contact with antigen presenting cells and becomes activated, T-cell receptor stimulates downstream signaling molecules to increase intracellular levels of calcium, which in turn activates calcineurin. Calcineurin is a serine/threonine phosphatase that de-phosphorylates the transcription factor, NF-AT, which in turn, translocated to the nucleus to transcribe pro-inflammatory genes, particularly IL-2 and related cytokines (Macian, 2005). Tacrolimus blocks this T-cell-mediated inflammatory circuit by inhibiting the function of calcineurin (Ganong, 2005). By reducing IL-2

production, T-cells are unable to differentiate into Th2 subtype, which consequently would prevent the activation of B-cells and IgE antibody secretion, thereby suppressing allergic reactions. Cyclosporin is another type of immunosuppressant medication that also acts as a calcineurin inhibitor. However, both tacrolimus and cyclosporin have posed risks of developing skin cancer or lymphoma and thus, not all individuals should be taking these medications.

In 2021, Tralokinumab (aka Adtralza) was approved in the United Kingdom and the United States as a therapeutic for atopic dermatitis (Freitas, et al., 2021). Tralokinumab is a human monoclonal antibody that specifically targets the cytokine, IL-13. Similar to IL-4, IL-13 is a pivotal cytokine involved in the generation of allergic diseases, such as AD via polarization of Th2 lymphocytes. There is increasing evidence that IL-13 is overexpressed in AD patients, and this is correlated with increased trafficking of inflammatory leukocytes and decreased function of epidermal barrier (Bieber, 2020). In parallel to these studies, there is a rationale for potentially targeting the effect of IL-4, the major cytokine that leads to Th2 response (Lernia, 2015). A combinatorial therapy of tralokinumab and IL-4 inhibitor may potentially synergize protection against the development and progression of AD. However, it is also important to consider the side effects because both IL-13 and IL-4 play vital roles in polarization of macrophages toward an M2, anti-inflammatory phenotype that regulate wound healing, tissue repair, and anti-parasitic infection (Liu, et al., 2021). By antagonizing the effects of these cytokines, there would be a systemic imbalance of pro-inflammatory vs. anti-inflammatory signatures, which can lead to many downstream tissue cytotoxicity.

5.3 Alternative Medicine

Traditional Chinese medicine (e.g., acupuncture, tai chi, and herbal products) has always been (and still is) a highly controversial topic in clinical studies and scientific reviews. This is because traditional Chinese medicine, along with many other types of oriental medicine, is not

backed by the scientific method. Recently however, many organizations such as the World Health Organization is beginning to recognize traditional Chinese medicine in its influential global medical compendium (Cyranoski, 2018). Some of these psychological and physical approaches used in oriental medicine can help alleviate pain conditions and improve quality of life (Matos, et al., 2021). In 2013, a study was conducted by Gu et. al to assess the effects of oral ingestion and topical application of Chinese herbal medicine (CHM) for the management and progression of atopic dermatitis. It was a randomized clinical trial consisting of 2306 participants, a mixture of both adults and children with atopic dermatitis. It was collectively shown that the participants in the CHM group (Both oral and topical administration) reported significantly less itching and higher quality of life (QoL) score than the placebo group (Gu, et al., 2013). Overall, CHM was well- tolerated in these clinical trials with no significant adverse side effects, except for a few mentions of reversible transaminitis (Goddard and Lio, 2015).

Another type of alternative medicine that is widely used in clinics is acupuncture and acupressure. Both are forms of traditional Chinese medicine that attempts to restore the body's Qi, which is defined as the vital force of energy that enables proper "flow of energy" (Taking Charge of Your Health & Wellbeing, 2022). Acupuncture requires a licensed, trained professional to insert hair-thin, sterile needles into different points of the body as an attempt to stimulate the body's meridians (i.e., energy pathways) and restore homeostasis (Clayton, 2022). On the other hand, acupressure is a non-invasive traditional Chinese medicine technique that also attempts to restore the body's Qi, but instead of needles, it utilizes gentle pressures to the skin. A study conducted in 2018 by Lee et. al showed that acupressure using a 1.2 mm acupellet at the L11 point significantly reduced the severity of itchiness and pruritus in AD patients compared to the placebo group (Lee, et al., 2012). Additionally, independent trials in animal models suggest that acupuncture and moxibustion (application of heat onto acupuncture points) can reduce AD- mediated pathology by decreasing the

expression of phosphorylated STAT6, which as described in the previous section, is the transcriptional factor that becomes activated in response to IL-4 signaling (Pfab et al., 2014). However, more research is needed to understand the specific mechanisms of acupuncture-mediated reduction in AD.

6. Conclusion

Atopic dermatitis and other hypersensitive immune disorders have been extensively investigated over the course of the last several decades, and they are still gaining much attention today. While this review article explored a relatively minor form of allergy, it is important to understand that allergic reactions can be deadly. For example, food allergy affects about 15 million people in the United States alone, with an average mortality around 150-200 individuals (Radke, et al., 2017). Some individuals contract anaphylactic shock against specific allergens, such as peanut butter, insect bites, or even certain medications. These shocks can be life-threatening if they don't immediately seek medical emergency (Cristol, 2020). Therefore, allergic diseases are not trivial, and more research is needed to identify biological pathways that are targeted for therapeutic purposes, while minimizing side effects.

Of course, there remains many important and unresolved questions about atopic dermatitis. One crucial issue is to properly distinguish atopic dermatitis from other forms of inflammatory conditions in the skin. The symptoms found in AD largely overlap with symptoms found in other skin allergic diseases, such as contact dermatitis, hay fever, and hives. While the Hanifin and Rajka criteria has been developed to clinically diagnose patients of AD, not all individuals have access to a clinical dermatologist. Therefore, many individuals end up consuming generic-brand medications that may alleviate temporary pain but don't fully treat the disease. While lifestyle changes can significantly reduce the development and progression of AD, it can often be burdensome for the individual to adopt these changes.

Altogether, the review paper examines the

characteristics of AD, starting with critical risk factors and diagnosis, followed by its molecular pathophysiology and current knowledge on therapeutic potential. Tackling some of these important unanswered questions regarding the pathophysiology of AD will undoubtedly advance the mechanistic understanding of allergy and hypersensitivity disorders, as well as help uncover novel therapeutic targets to combat the progression of this disease.

Acknowledgment

I gratefully thank and acknowledge my mentor, Min Sub Lee for his guidance, advice, and continuous support throughout the process of my article. This wouldn't have been possible without his exchange of input on technical information, research, and shared knowledge. I would also like to thank Katia Chadaideh, PhD and her lectures on allergies and autoimmune disorders during her summer course, Introduction to Evolutionary Medicine, Harvard Precollege Summer Program.

References

- Akan, A., et. al., (2020). Diagnosis of atopic dermatitis in children: comparison of the Hanifin-Rajka and the United Kingdom Working Party criteria. *Allergologia et immunopathologia*, 48(2), 175–181. <https://doi.org/10.1016/j.aller.2019.07.008>
- Amin K. (2012). The role of mast cells in allergic inflammation. *Respiratory medicine*, 106(1), 9–14. <https://doi.org/10.1016/j.rmed.2011.09.007>
- Arthur, J. C., et. al., (2010). Cutting edge: NLRP12 controls dendritic and myeloid cell migration to affect contact hypersensitivity. *Journal of immunology (Baltimore, Md. : 1950)*, 185(8), 4515–4519. <https://doi.org/10.4049/jimmunol.1002227>
- Berger A. (2000). Th1 and Th2 responses: what are they?. *BMJ (Clinical research ed.)*, 321(7258), 424. <https://doi.org/10.1136/bmj.321.7258.424>
- Bieber T. (2020). Interleukin-13: Targeting an underestimated cytokine in atopic dermatitis. *Allergy*, 75(1), 54–62. <https://doi.org/10.1111/all.13954>
- Brandt, E. B., & Sivaprasad, U. (2011). Th2 Cytokines and Atopic Dermatitis. *Journal of clinical & cellular immunology*, 2(3), 110. <https://doi.org/10.4172/2155-9899.1000110>
- Bufford, J. D., & Gern, J. E. (2005). The hygiene hypothesis revisited. *Immunology and allergy clinics of North America*, 25(2), 247–vi. <https://doi.org/10.1016/j.iac.2005.03.005>
- Carmela Avena-Woods, B. S. P. (n.d.). *Overview of atopic dermatitis*. AJMC. Retrieved September 1, 2022, from <https://www.ajmc.com/view/overview-of-atopic-dermatitis-article>
- Clayton, V. (2022, August 8). *Acupuncture vs. acupressure: What's the difference?* Forbes. Retrieved September 1, 2022, from <https://www.forbes.com/health/body/acupuncture-vs-acupressure/>
- Cristol, H. (2020). *Anaphylactic shock: What you should know*. WebMD. Retrieved September 1, 2022, from <https://www.webmd.com/allergies/anaphylactic-shock-facts#:~:text=Anaphylactic%20shock%20is%20a%20rare,911%20for%20emergency%20medical%20help.>
- Cyranoski D. (2018). Why Chinese medicine is heading for clinics around the world. *Nature*, 561(7724), 448–450. <https://doi.org/10.1038/d41586-018-06782-7>
- Di Lernia V (January 2015). "Therapeutic strategies in extrinsic atopic dermatitis: focus on inhibition of IL- 4 as a new pharmacological approach". *Expert Opinion on Therapeutic Targets*. 19 (1): 87–96.
- Engelbrechtsen, K. A., et. al., (2016). The effect of environmental humidity and temperature on skin barrier function and dermatitis. *Journal of the European Academy of Dermatology and Venereology : JEADV*, 30(2), 223–249. <https://doi.org/10.1111/jdv.13301>
- FDA. (2018). *Asthma: The hygiene hypothesis*. U.S. Food and Drug Administration. Retrieved September 1, 2022, from <https://www.fda.gov/vaccines-blood-biologics/consumers-biologics/asthma-hygiene-hypothesis>

- Flohr, C., & Mann, J. (2014). New insights into the epidemiology of childhood atopic dermatitis. *Allergy*, 69(1), 3–16. <https://doi.org/10.1111/all.12270>
- Freitas, E., Guttman-Yassky, E., & Torres, T. (2021). Tralokinumab for the Treatment of Atopic Dermatitis. *American journal of clinical dermatology*, 22(5), 625–638. <https://doi.org/10.1007/s40257-021-00613-8>
- Goddard, A. L., & Lio, P. A. (2015). Alternative, Complementary, and Forgotten Remedies for Atopic Dermatitis. *Evidence-based complementary and alternative medicine : eCAM*, 2015, 676897. <https://doi.org/10.1155/2015/676897>
- Goldman, R. (2017, October 12). *How much weight can you lose in a month?* Healthline. Retrieved September 1, 2022, from <https://www.healthline.com/health/food-nutrition/weight-loss-in-a-month>
- Gu, S., Yang, A. W., Xue, C. C., et al., (2013). Chinese herbal medicine for atopic eczema. *The Cochrane database of systematic reviews*, (9), CD008642. <https://doi.org/10.1002/14651858.CD008642.pub2>
- Hahn, E. L., & Bacharier, L. B. (2005). The atopic march: the pattern of allergic disease development in childhood. *Immunology and allergy clinics of North America*, 25(2), 231–v. <https://doi.org/10.1016/j.iac.2005.02.004>
- Harrison, D. L., Fang, Y., & Huang, J. (2019). T-Cell Mechanobiology: Force Sensation, Potentiation, and Translation. *Frontiers in physics*, 7, 45. <https://doi.org/10.3389/fphy.2019.00045>
- Hoffman, W., Lakkis, F. G., & Chalasani, G. (2016). B Cells, Antibodies, and More. *Clinical journal of the American Society of Nephrology : CJASN*, 11(1), 137–154. <https://doi.org/10.2215/CJN.09430915>
- Imai Y. (2019). Interleukin-33 in atopic dermatitis. *Journal of dermatological science*, 96(1), 2–7. <https://doi.org/10.1016/j.jdermsci.2019.08.006>
- Indra A. K. (2013). Epidermal TSLP: a trigger factor for pathogenesis of atopic dermatitis. *Expert review of proteomics*, 10(4), 309–311. <https://doi.org/10.1586/14789450.2013.814881>
- Jacobs, S. R., & Damania, B. (2012). NLRs, inflammasomes, and viral infection. *Journal of leukocyte biology*, 92(3), 469–477. <https://doi.org/10.1189/jlb.0312132>
- Jennemann, R., et al., (2007). Integrity and barrier function of the epidermis critically depend on glucosylceramide synthesis. *The Journal of biological chemistry*, 282(5), 3083–3094.
- Kane SP. Hydrocortisone, ClinCalc DrugStats Database, Version 2022.08. ClinCalc: <https://clincalc.com/DrugStats/Drugs/Hydrocortisone> Updated August 24, 2022. Accessed September 1, 2022.
- Kane SP. The Top 300 of 2020, ClinCalc DrugStats Database, Version 2022.08. ClinCalc: <https://clincalc.com/DrugStats/Top300Drugs.aspx>. Updated August 24, 2022. Accessed September 1, 2022.
- Kondělková, K., et al., (2010). Regulatory T cells (TREG) and their roles in immune system with respect to immunopathological disorders. *Acta medica (Hradec Kralove)*, 53(2), 73–77. <https://doi.org/10.14712/18059694.2016.63>
- Kuzume, K., & Kusu, M. (2007). Before-birth climatologic data may play a role in the development of allergies in infants. *Pediatric allergy and immunology : official publication of the European Society of Pediatric Allergy and Immunology*, 18(4), 281–287. <https://doi.org/10.1111/j.1399-3038.2006.00526.x>
- Lee, K. C., Keyes, A., Hensley, J. R., et al., (2012). Effectiveness of acupressure on pruritus and lichenification associated with atopic dermatitis: a pilot trial. *Acupuncture in medicine : journal of the British Medical Acupuncture Society*, 30(1), 8–11. <https://doi.org/10.1136/acupmed-2011-010088>
- Lio P. A. (2013). Non-pharmacologic therapies for atopic dermatitis. *Current allergy and asthma reports*, 13(5), 528–538. <https://doi.org/10.1007/s11882-013-0371-y>
- Liu, J., et al. New insights into M1/M2 macrophages: key modulators in cancer progression. *Cancer Cell Int* 21, 389 (2021). <https://doi.org/10.1186/s12935-021-02089-2>

- Macian F. (2005). NFAT proteins: key regulators of T-cell development and function. *Nature reviews. Immunology*, 5(6), 472–484. <https://doi.org/10.1038/nri1632>
- Matos, L. C., Machado, J. P., Monteiro, F. J., & Greten, H. J. (2021). Understanding Traditional Chinese Medicine Therapeutics: An Overview of the Basics and Clinical Applications. *Healthcare (Basel, Switzerland)*, 9(3), 257. <https://doi.org/10.3390/healthcare9030257>
- Mayo Foundation for Medical Education and Research. (2022, August 1). *Tacrolimus (oral route) side effects*. Mayo Clinic. Retrieved September 1, 2022, from <https://www.mayoclinic.org/drugs-supplements/tacrolimus-oral-route/side-effects/drg-20068314?p=1#:~:text=Tacrolimus%20is%20used%20together%20with,medicines%20known%20as%20immunosuppressive%20agents>
- McLean W. H. (2011). The allergy gene: how a mutation in a skin protein revealed a link between eczema and asthma. *F1000 medicine reports*, 3, 2. <https://doi.org/10.3410/M3-2>
- Meyer-Hoffert, U., & Schröder, J. M. (2011). Epidermal proteases in the pathogenesis of rosacea. *The journal of investigative dermatology. Symposium proceedings*, 15(1), 16–23. <https://doi.org/10.1038/jidsymp.2011.2>
- Minniti, F., et al., (2014). Breast-milk characteristics protecting against allergy. *Endocrine, metabolic & immune disorders drug targets*, 14(1), 9–15. <https://doi.org/10.2174/1871530314666140121145045>
- Miyake, Y., Sasaki, S., Tanaka, K., & Hirota, Y. (2010). Dairy food, calcium and vitamin D intake in pregnancy, and wheeze and eczema in infants. *The European respiratory journal*, 35(6), 1228–1234. <https://doi.org/10.1183/09031936.00100609>
- Mogensen T. H. (2009). Pathogen recognition and inflammatory signaling in innate immune defenses. *Clinical microbiology reviews*, 22(2), 240–273. <https://doi.org/10.1128/CMR.00046-08>
- Nall, R. (2020, August 13). *Contact dermatitis: Causes, symptoms, and treatment*. Healthline. Retrieved September 1, 2022, from <https://www.healthline.com/health/contact-dermatitis>
- National Center for Biotechnology Information (2022). PubChem Compound Summary for CID 5754, Hydrocortisone. Retrieved September 2, 2022 from <https://pubchem.ncbi.nlm.nih.gov/compound/Hydrocortisone>.
- NENA Skincare. (2020, February 13). *Mineral water benefits for skin and health*. NENA Skincare. Retrieved September 1, 2022, from <https://us.nenaskincare.com/blogs/wellness/mineral-water-benefits-for-skin-and-health>
- Okada, H., Kuhn, C., Feillet, H., & Bach, J. F. (2010). The 'hygiene hypothesis' for autoimmune and allergic diseases: an update. *Clinical and experimental immunology*, 160(1), 1–9. <https://doi.org/10.1111/j.1365-2249.2010.04139.x>
- Pelucchi, C., et al., (2013). Pet exposure and risk of atopic dermatitis at the pediatric age: a meta-analysis of birth cohort studies. *The Journal of allergy and clinical immunology*, 132(3), 616–622.e7. <https://doi.org/10.1016/j.jaci.2013.04.009>
- Pfab, F., et al., (2014). Acupuncture for allergic disease therapy--the current state of evidence. *Expert review of clinical immunology*, 10(7), 831–841. <https://doi.org/10.1586/1744666X.2014.924855>
- Radke, T. J., et al., (2017). Restaurant Food Allergy Practices - Six Selected Sites, United States, 2014. *MMWR. Morbidity and mortality weekly report*, 66(15), 404–407. <https://doi.org/10.15585/mmwr.mm6615a2>
- Raetz, C. R., & Whitfield, C. (2002). Lipopolysaccharide endotoxins. *Annual review of biochemistry*, 71, 635–700. <https://doi.org/10.1146/annurev.biochem.71.110601.135414>
- Sandilands, A., Sutherland, C., Irvine, A. D., & McLean, W. H. (2009). Filaggrin in the frontline: role in skin barrier function and disease. *Journal of cell science*, 122(Pt 9), 1285–1294. <https://doi.org/10.1242/jcs.033969>
- Schlapbach, C., & Simon, D. (2014). Update on skin allergy. *Allergy*, 69(12), 1571–1581. <https://doi.org/10.1111/all.12529>
- Shi, V. (n.d.). *Common air pollutants that trigger*

eczema | *learnskin*. Retrieved September 2, 2022, from <https://www.learnskin.com/articles/how-does-air-pollution-affect-eczema>

Siegfried EC, Hebert AA. Diagnosis of Atopic Dermatitis: Mimics, Overlaps, and Complications. *J Clin Med*. 2015 May 6;4(5):884-917. doi: 10.3390/jcm4050884. PMID: 26239454; PMCID: PMC4470205.

Silk properties. marimichel.nl. (2022). Retrieved September 1, 2022, from <https://marimichel.nl/pages/silk-properties>

Soumelis, V., et al., (2002). Human epithelial cells trigger dendritic cell mediated allergic inflammation by producing TSLP. *Nature immunology*, 3(7), 673–680. <https://doi.org/10.1038/ni805>

Stone, K. D., Prussin, C., & Metcalfe, D. D. (2010). IgE, mast cells, basophils, and eosinophils. *The Journal of allergy and clinical immunology*, 125(2 Suppl 2), S73–S80. <https://doi.org/10.1016/j.jaci.2009.11.017>

Tee-Melegrito, R. A. (2021). *What is the difference between atopic and contact dermatitis?* Medical News Today. Retrieved September 1, 2022, from <https://www.medicalnewstoday.com/articles/atopic-dermatitis-vs-contact-dermatitis#differences>

Théry, C., & Amigorena, S. (2001). The cell biology of antigen presentation in dendritic cells. *Current opinion in immunology*, 13(1), 45–51. [https://doi.org/10.1016/s0952-7915\(00\)00180-1](https://doi.org/10.1016/s0952-7915(00)00180-1)

Trikamjee, T., et al., (2021). Nutritional Factors in the Prevention of Atopic Dermatitis in Children. *Frontiers in pediatrics*, 8, 577413. <https://doi.org/10.3389/fped.2020.577413>

Tsang, M. S., Hou, T., Chan, B. C., & Wong, C. K. (2021). Immunological Roles of NLR in Allergic Diseases and Its Underlying Mechanisms. *International journal of molecular sciences*, 22(4),

1507. <https://doi.org/10.3390/ijms22041507>

U.S. Department of Health and Human Services. (2022, June 22). *Atopic dermatitis*. National Institute of Arthritis and Musculoskeletal and Skin Diseases. Retrieved September 1, 2022, from <https://www.niams.nih.gov/health-topics/atopic-dermatitis>

Varothai, S., Nitayavardhana, S., & Kulthanan, K. (2013). Moisturizers for patients with atopic dermatitis. *Asian Pacific journal of allergy and immunology*, 31(2), 91–98.

Watson, S. (2021). *What is atopy?* WebMD. Retrieved September 1, 2022, from <https://www.webmd.com/skin-problems-and-treatments/eczema/what-is-atopy>

WebMD. (2022). *Eczema: Types, symptoms, causes, diagnosis, and treatment*. WebMD. Retrieved September 1, 2022, from <https://www.webmd.com/skin-problems-and-treatments/eczema/atopicdermatitis-eczema>

What is Qi? (and other concepts). Taking Charge of Your Health & Wellbeing. (2022). Retrieved September 1, 2022, from <https://www.takingcharge.csh.umn.edu/explore-healing-practices/traditional-chinese-medicine/what-qi-and-other-concepts>

William F. Ganong (8 March 2005). *Review of medical physiology* (22nd ed.). Lange medical books. p. 530. ISBN 978-0-07-144040-0.

Yang, G., et al., (2020). Skin Barrier Abnormalities and Immune Dysfunction in Atopic Dermatitis. *International journal of molecular sciences*, 21(8), 2867. <https://doi.org/10.3390/ijms21082867>

Zhong, Y., Kinio, A., & Saleh, M. (2013). Functions of NOD-Like Receptors in Human Diseases. *Frontiers in immunology*, 4, 333. <https://doi.org/10.3389/fimmu.2013.00333>

COVID-19: Survey on Situation, Enthusiasm and Attitude for Vaccination of SARS-CoV-2 Boosters

Ming-Liu Li^{1*}, Ya-Dan Huang¹

¹Guangdong Country Garden School, Foshan, Guangdong province China

Received March 16, 2022; Revised February 14, 2023; Accepted, March 10, 2023

Abstract

COVID-19 has become a global phenomenon and a top priority for the world, making vaccine implementation a key tool for countries to address the problem. Boost refers to one additional supplementary vaccination following the completion of two doses of the SARS-CoV-2 vaccine, resulting in more than a 10-fold increase in antibody levels. The SARS-CoV-2 vaccine can maintain human immunity to the virus after antibodies have waned for some time. This study conducted a survey of the prevalence of SARS-CoV-2 vaccination in Guangdong Province, investigating SARS-CoV-2 vaccination among students and some educators, knowledge of the vaccine, attitudes, and the status of booster needle vaccination. Excel and SPSS software were used to analyze 166 samples. Five hypotheses were tested according to the data analysis, which are about the vaccination situation of SARS-CoV-2 vaccine, the relationship between people's understanding of the vaccine and the vaccination rate, people's attitude and enthusiasm towards the vaccine and the side effects of the booster injection of SARS-CoV-2 vaccine. This survey suggests that Policy-influenced vaccination intentions, rather than personal knowledge and understanding of the vaccine, were more strongly associated with the vaccination rates. , providing better reference to other pandemic and to areas with low vaccination coverage.

Keywords: SARS-CoV-2 vaccine, Questionnaires, Booster, Vaccination Policy

1. Introduction

Up to December 22th, 2022, according to incomplete statistics, a total of 641.93 million people has been diagnosed globally, and a total of 6.64 million people have died from COVID-19 infection (World Epidemic Statistics, 2023). The virus causing the symptoms of covid-19 spreads extremely rapidly, and there is currently no specific cure for this condition. Therefore, the development of vaccines that can prevent infection with this virus is important. Vaccines are autoimmune preparations of pathogenic microorganisms (e.g., bacteria, rickettsia, viruses, etc.) and their metabolites that have been artificially

attenuated, inactivated, or genetically modified to prevent infectious diseases (Vaccine, 2022). The advent of vaccines has allowed people to fundamentally avoid infections with a number of diseases that they do not need to struggle with the pain caused by these viruses or to think with great thought about how to treat them.

For decades, scientific research teams around the world have had a model of vaccine development to guarantee that vaccine development is safe and effective. However, pandemic diseases similar to COVID-19 have extremely high demands for the speed and efficiency of vaccine development, with no exception the need for a vaccine that can protect

* Corresponding Author
krischellelml_20xx@foxmail.com

Advisor: Ya-dan Huang
0024479@brightscholar.com

against the SARS-CoV-2 virus in countries during an outbreak. So different patterns of vaccine development may be followed during a pandemic and production levels may need to be expanded at an early age. There are already some vaccines that have completed the initial technological development and gradually move to the clinical stage, heralding the future with more and more vaccines coming to market. In 2021, WHO urgently approved the Kochi vaccination. Already in mid-April 2021, about 180 vaccine candidates were in the clinical stage, and another 88 were still in the clinical stage (Nagy and Alhatlani, 2021). And three vaccines, Pfizer / BioNTech, Moderna mRNA vaccines and Johnson viral vector vaccines, and AstraZeneca viral vector vaccine, Oxford, UK, are already in emergency use (Nagy and Alhatlani, 2021).

Today, these vaccines are at the forefront of innovative technologies. Viral vectors and nucleic acid vaccines not only are highly versatile technologies but can expedite vaccine production and scale-up production, meeting the demand for a vaccine among people during a pandemic (Wang, et al., 2020).

The technology of live attenuated and inactivated vaccines is more mature and can provide effective antibody immunity, but their development cycle is long, and the cost is high. The development of gene vaccines (mRNA vaccines, DNA vaccines) that require less time and are less expensive to produce in recent years has provided a good response to the sudden onset of disease but are less immune. Apart from that, adenovirus vectored vaccine is also progressing, and although the maturity of the industry is not high, the faster development rate, the lower cost, and the particularly good immune effect (both humoral and cellular immunity are particularly good) of this vaccine and can exert adjuvant effect are also very promising. These vaccines of different manufacturing processes are more or less applied to prevent a wide range of diseases. For instance, Human papilloma virus vaccine as the vaccine against tumor prevention. It helped many people, especially women, to avoid a series of diseases and cancers caused by HPV (HPV) virus (Maryam, et al., 2018).

The current development of vaccines against such

viruses has advanced, driven by societal needs as well as the efforts of researchers. However, the study still faces many problems, technical advances in bioinformatics and molecular approaches are needed to help recognizing novel epitopes, which leading to better vaccine design. And a deeper understanding of the SARS-CoV-2 virus, the development of different vaccines, and better vaccine design will all require time to study.

Booster needle refers to a vaccine dose that preserves human immunity to the virus in a supplemental manner, based on antibody waning, after completion of vaccination (COVID-19 Booster).

The significance of needle strengthening is particularly relevant, since a decline in antibody levels after a period of vaccination can lead to an increased probability that an individual will become ill.

For some people with low immunity, it is important to strengthen the vaccination of SARS-CoV-2 vaccine. Because even if they were given two shots of vaccine, they would not have enough immune response. In a study by Kamar et al., follow-up studies were conducted on patients with organ transplants with low immune function, and antibody levels increased by nearly thirty percent after re-administration of booster injections (2021). This study illustrates the vaccination of booster needles, especially for these people, is reasonable to be required and proved to be effective.

On 12 August 2021, the US Food and Drug Administration revised the emergency use authorization (EUAs) for the Pfizer biotechnology COVID-19 vaccine and Moderna COVID-19 vaccine to allow SARS-CoV-2 vaccines use in specific immunocompromised individuals who have undergone solid organ transplantation or are diagnosed with additional doses that are considered to have an equivalent level of immune compromise (Shekhar, et al., 2021). However, to safeguard the health of vaccinees and avoid the potential health risks associated with multiple vaccinations, the CDC claims that the third dose should be given three weeks after the second dose of Pfizer vaccine and the Moderna vaccine four weeks later. In addition, COVID-19 vaccination for adults 65 years of age and older, living in a nursing facility for an extended

period of time, with an underlying medical problem, or working or living in a high-risk environment. For example, individuals at increased risk of exposure to and transmission of COVID-19 due to occupational or institutional settings, such as first responders, educators, food and agriculture workers, correctional personnel, and public transport workers, are eligible for a dose of COVID-19 vaccine (Shekhar, et al., 2021).

Vaccination with the inactivated vaccine accounts for a large proportion in China, whereas antibody levels in humans decline substantially after several months of vaccination with the inactivated vaccine, and vaccination with the SARS-CoV-2 vaccine booster needle is particularly important in China (New crown vaccine third needle, officially opened!, 2021). On 17, Sept 2021 vaccination with booster needle was officially opened and the vaccinated population under 12 years old will be included in the scope of vaccination. (New crown vaccine third needle, officially opened!, 2021).

2. Materials and Methods

2.1 Data Collection

The distribution and collection of original data for this 166 questionnaires were carried out through the Wenjuanxing (<https://www.wjx.cn/>), a typic website designed for managing questionnaires, for survey conduction. No patients or members of the public were directly involved in this study as no primary data were collected.

2.2 Statistical Analysis

The collected data of questionnaire were exported into Excel table, and analyzed using the online application software SPSSAU (version 23.0). A total of seventeen questions were catergorised into five

dimensions i.e personal information, vaccination status of the SARS-CoV-2 vaccine, personal knowledge about the vaccine, personal attitudes towards vaccination, and a follow-up survey of SARS-CoV-2 vaccine booster. The effectiveness and fitness of the research items of the questionnaire were studied using an item-analysis (table 2 and 5). The Pearson correlation coefficient(r) (table 6) and Linear Regression model (table 7) was used to determine whether there was statistical evidence to indaicate the relationship between the vaccination rate and personal knowledge and understanding of participants. The set scale of questions at same dimension were compared using t-test and the *p*-values were then recorded in the table where significances were denoted with an asterisk.

3. Results and Discussion

3.1 Personal Information

The three questions at the outset is about the subjects' age, residence and range of activity, and occupation. According to these questions, we can preliminarily determine the population of the sample and analyze its frequency. Ages 12to 18 accounted for the highest 70.48%. There is relatively more "Foshan" in the sample, with a proportion of 56.02%. In addition, more than 70% of the sample were "student". There is also 13.25% of samples are educators, which means the samples are mainly consists of dense contact population in educational organizations in Foshan and other districts of Guangdong Province, most of them are teenagers. The intensive ditribution of participants' occupation and age will give more clues to the the association between personal attitude and willing to take the SARS-CoV-2 vacine among the well-eduacted area.

3.2 Vaccination Status

Hypothesis 1: people who have completed the second vaccination accounts for the heaviest proportion in this section

Table 1. questions and options of vaccination status of the SARS-CoV-2 vaccine with scale

	Classification	Question	Options	scale	Skip Logic	
4	single choice	Which of the following is your current SARS-CoV-2 vaccination status?	A	I have not started vaccination yet	0	to Q9
			B	the first injection has been completed	1	to Q10
			C	the second injection has been completed	2	to Q11
			D	the third injection has been completed	3	to Q12

Table 2. the result of item-analysis of question 4

Question	Groups (mean ± SD)		t (CR)	p
	high proficiency group (n=45)	low proficiency group (n=46)		
4	2.11±0.44	2.37±0.49	2.66	0.009**

** $p < 0.01$.

As shown in question 4(table 1), there is a scale from 0 to 3 for question 4. The item-analysis was performed to test whether the research items of the questionnaire are effective and appropriate (SPSS Online, 2022). The data was summed and divided into high score and low score groups (bounded by 27% and 73% quantiles), and then t-test was used to compare the differences between the averages of high score and low score groups (table 2).

The item-analysis of question 4 based on the participants' vaccination status showed a significance difference between high proficiency group and low proficiency group ($p < 0.01$) (table 2), indicating the scale was reasonably designed. We found that the second injection has been completed among this sampling accounted for the highest 73.49%. According to China's policy, for the types of people who often gather, it is required to complete two doses of SARS-CoV-2 vaccination, and centralized vaccination sites and vaccination time have been set up in many schools. And we can see most people tend to follow the encouragement of policies, which probably relates with the fact that nearly 90% participants are in education system.

However, despite the government opening and encouraging vaccination with the booster doses in the second half of 2021, vaccination rates with boosters remained low. Even though the primary population surveyed for this time was students within Guangdong Province, the number vaccinated with a boosters of SARS-CoV-2 was only 1.81% of the total. According to the province's news online, the boosters is mainly aimed at people at high risk of input, such as workers from customs, border checks, aviation, isolation sites, site directed medical facilities, people with relatively low immune function, people over 60 years of age, and people who are going to areas or countries at high risk of outbreaks because of the need for work, learning, and communication (New crown vaccine third needle,

officially opened!, 2021). And most of the subjects surveyed did not belong to one of the above three points, thus also explaining the lower vaccination rate. However, the school is a very densely contacted place, and even though the students surveyed are mostly teenagers and have better immune function, but the students out of the school have numerous staff members ranging in age from 18 to retirement and are all social, exposure to pathogens is not without it, Nor is there (although not much) incentive for low functioning individuals among the students. According to the present findings, the greatest beneficiaries of booster needles are immunocompromised populations (Shekhar, et al., 2021), but the decline in antibody numbers is seen some time after vaccination for all, and thus whether and how vaccination might be encouraged for students and, similarly, for populations prone to close contacts might require further research to prove and improve.

Based on this data, this survey examined the level of knowledge about the vaccine and attitudes. And to seek correlations and effects of these two aspects on the overall vaccination situation.

3.3 Personal Knowledge of Vaccine

Hypothesis 2: there is a relatively strong and positive correlation between the vaccination rate and people's understanding of the vaccine.

The purpose of this part of the questionnaire design was to see, by 3-question questions, the level of knowledge that the surveyed people had about vaccination policy and about the basics of vaccines in China. This section designs a 1 to 5 scale to measure how well people know about the vaccine(table 3). Of these, 1 represented that person didn't have any knowledge of the vaccine, 2 represented that person had little knowledge of the vaccine, 3 represented

that person had some knowledge of the vaccine but was not comprehensive, 4 represented that person had

much knowledge of the vaccine, and 5 represented that person had much knowledge of the vaccine.

Table 3. questions and options of the knowledge of SARS-CoV-2 vaccine with scale

	Classification	Question	Options		Skip Logic	Scale
5	single choice	Do you know that the SARS-CoV-2 vaccine can be vaccinated in China now?	A	yes	no	5
			B	no		1
6	single choice	Do you know how many shots are required for the SARS-CoV-2 vaccine in China?	A	one needle	no	3
			B	two needles		5
			C	three needles		3
			D	I do not know		1
			E	any number of needles are OK		1

Question 4 listed in table 4 is to check the knowledge of the investigated people about the vaccine (principle of action, group immunity). The scores corresponding to each option are shown in the figure. The text in the column ‘Options’ is the situation when the original data is collected. After collecting the original data, these data are processed and the text options are converted into the corresponding numbers of the scale. There are six

options in this question, of which three are the wrong description of the vaccine and the other three are the correct description of the vaccine. The wrong score is - 1 and the correct score is 1. This question is a multiple-choice question. In data processing, the total score of a single sample is converted into a scale of 1-5 points. The score conversion scale mechanism of question 7 is as follows: 0 for - 3, 1 for - 2 and - 1, 2 for 0, 3 for 1, 4 for 2, and 5

Table 5. the result of item-analysis of question 5, 6, and 7

Item-analysis Results				
Question	Groups (mean ± SD)		t (CR)	p
	high proficiency group (n=45)	low proficiency group (n=46)		
5	5.00±0.00	5.00±0.00	null	null
6	3.36±1.30	3.43±1.26	0.295	0.768
7	3.96±1.09	4.39±0.74	2.227	0.029*

It was assumed that the differences between 166 questionnaires are a result of different, yet related, intervention effects being estimated. As shown in table 6, the correlation analysis based on Pearson correlation coefficient was used to study the correlation between vaccination rate and people's understanding of vaccination policy and vaccine knowledge (SPSS Online, 2022). The specific analysis showed that the correlation coefficient between question 4 and question 6 is -0.148, close to

0, and the *p*-value is 0.056. The correlation coefficient between question 4 and question 7 is 0.037, close to 0, and the *p*-value is 0.637 (table 6). The result showed that there is no direct correlation between the vaccination rate and people's understanding of the vaccine. This may be due to the fact that the sample distribution is not strictly normal. And it's hard to assess precisely the willing and understanding about the vaccination with limited questions.

Table 6. the result of Pearson correlation of question 6 and 7 with question 4

Pearson Correlation (Detail)		
Question4		
Question 6	Coefficient	-0.148
	<i>p</i>	0.056
Question 7	Coefficient	0.037
	<i>p</i>	0.637

Regression analysis was then used to explore the influence relationship between X (quantitative or categorical) and Y (quantitative), whether there is an influence relationship, and how about the influence direction and degree (The SPSSAU project, 2021) (table 7).

Table 7. the result of Linear Regression of question 5, 6 and 7 with question

Parameter Estimates (n=166)									
	Unstandardized Coefficients			<i>t</i>	<i>p</i>	VIF	<i>R</i> ²	Adj <i>R</i> ²	<i>F</i>
	<i>B</i>	Std. Error	<i>Beta</i>						
Constant	2.267	0.712	-	3.183	0.002**	-	0.023	0.005	F (3,162) =1.269, P=0.287
Question 5	0.013	0.136	0.007	0.095	0.925	1.021			
Question 6	-0.067	0.036	-0.146	-1.872	0.063	1.012			
Question 7	0.018	0.048	0.029	0.37	0.712	1.013			
Dependent Variable: Question 4 (vaccination rate)									
D-W: 0.995									

** *p*<0.01.

By setting questions 5, 6, 7 as x and 4 as y, this survey explores whether people’s understanding of the vaccine is related to the vaccination rate. According to Table 11, taking question 6 and question 7 as independent variables and question 4 as dependent variables for linear regression analysis. From the above table, we can see that the model formula is Question 4 = 2.267 + 0.013 * question 5 - 0.067 * question 6 + 0.018 * question 7, and the R square value of the model is 0.023, which means that people’s understanding of vaccination policies and knowledge can explain the 2.3% change of vaccination rate. It is found that the model does not pass the F test (F = 1.269, p = 0.287), which means that people’s understanding of vaccination policies and knowledge will not have a significant impact on the vaccination rate, so it is impossible to specifically analyze the impact of independent variables on dependent variables.

According to the data analysis results of this part, hypothesis 2 is not tenable. In a word, people’s knowledge of the vaccine can slightly explain for vaccination, but it cannot directly affect the vaccination rate in this sampling.

3. 4 Attitude to Vaccination

Hypothesis 3: The willingness of taking SARS-CoV-2 vaccine booster are more determined by the subjective wishes of the people, whereas the number of previous doses that people have taken are more determined by objective conditions.

In this section, four questions are designed for investigate people's attitudes towards vaccines (table 8). For those who did not complete the three doses of vaccine, the reasons why they did not complete the vaccination were collected, and the respondents were asked about their motivation and reasons for vaccination.

For those who had not started the SARS-CoV -2 vaccine, the reasons were only physical and the organization did not have the vaccination (question 8), indicating that those who did not start the vaccination were all vaccinated without any SARS-CoV-2 because of objective conditions, whereas one chose 'no formal place to go for vaccination', illustrating that some regions also need to be investigated for formal vaccination points.

Table 8. questions and options of question 8, 9, 10 and 11

	Classification	Question	Options	
8	multiple choice	Which of the following is the reason why you did not start getting the SARS-CoV-2 vaccine?	A	physical reasons (such as inability to vaccinate due to taking certain drugs, history of drug allergy, etc.)
			B	age reason
			C	do not have time to vaccinate
			D	I do not think it is necessary to vaccinate
			E	there is no formal organization in the community that can be vaccinated
			F	I think the vaccine is unreliable
9	multiple choice	Which of the following is the reason why you did not get the second shot of the new coronavirus vaccine?	A	after the completion of the first injection, the time for the second injection has not been reached
			B	physical reasons (such as inability to vaccinate due to taking certain drugs, history of drug allergy, etc.)
			C	age reason
			D	do not have time to vaccinate
			E	I do not think it is necessary to vaccinate
			F	there is no formal organization in the community that can be vaccinated
			G	I think the vaccine is unreliable
10	multiple choice	Which of the following is the reason why you did not get the third shot of the new coronavirus vaccine?	A	after the completion of the second injection the time for the second injection has not been reached
			B	physical reasons (such as inability to vaccinate due to taking certain drugs, history of drug allergy, etc.)
			C	age reason
			D	do not have time to vaccinate
			E	I do not think it is necessary to vaccinate
			F	there is no formal organization in the community that can be vaccinated
			G	I think the vaccine is unreliable
11	multiple choice	What is your reason / motivation for getting the booster of SARS-CoV-2 Vaccine	A	I'm required to vaccinate (such as occupational requirements or work area requirements)

For question 9, there is 33.33% people choose option “After the completion of the first injection, the time for the second injection has not been reached”, “Physical reasons (such as being unable to vaccinate due to taking certain drugs, having a history of drug allergy, etc.)”, and “There is no formal organization

in the community that can be vaccinated”. Which are all objectives reasons.

For question 10, there is 53.28% people choose option “After the completion of the second injection, the time for the third injection has not been reached” and 24.59% people choose option “Age reason”,

which are both objective. There does have a little proportion of subjective reason that people did not get the third shot of SARS- CoV-2 vaccine, but most of them is because time has not been reached or because of age.

For the conclusion above, an addition question was asked to those people who has not injected the booster: if your physical reasons are excluded, are you willing to complete the three injections of the vaccine in time? 89.84% of them say “yes”.

According to question 11, there is 79.75% people think vaccination is very important, because it can prevent the epidemic starts from them, which shows people are already well equipped with a sense of preparedness as well as citizenship, and the vast

majority are vaccinated because of a subjective willingness to vaccinate with booster needles. Further, a large proportion of those vaccinated with the booster needle were determined by subjective will.

In summary, the hypothesis 3 holds according to the data analysis in this section.

Hypothesis 4: Most people have a negative attitude towards universal vaccination.

Two questions were designed to see people’s attitudes towards herd immunity (table 9). Both issues had acquisition of data for all samples.

Table 9. questions and options of question 12 and 13.

	Classification	Question	Options	
12	single choice	Do you think we should have universal vaccination?	A	It should be noted that the higher the vaccination rate, the more likely it is to achieve herd immunity, except that physical conditions do not allow it.
			B	More respect for personal wishes and ideas
			C	It is not necessary at all
13	single choice	What do you think of the long- term coexistence with COVID- 19 without vaccination?	A	Disagree. If everyone thinks so, there will be no herd immunity
			B	I do not quite agree. Once you come into contact with the virus, you will be infected
			C	Quite agree. I think this is a good idea and saves a lot of things
			D	not always. It should be discussed in different regions. There should be no big problem if we do not vaccinate in places that are not epidemic areas

From question 12, there were relatively more people choosing a with a proportion of 78.31%. It is believed that universal vaccination is very important. From question 13 (Table 18), more than 3% of the sample had a selected, and 37.35% had B. Neither option agreed on the way vaccination did not work with the COVID-19 outbreak over a long period of time.

According to the frequency analysis above, the hypothesis 4 does not hold. The vast majority were positive for herd immunity and universal vaccination. Moreover, the determinants of attitude towards

vaccination of people in fact are complex and reciprocal.

It is inferred from the test results of the hypothesis, and the conclusions drawn, that one is willing to vaccinate whether the vaccination is driven by each other for subjective and objective reasons, that is, endo - and exo driving. And the two drives differ in the stage at which they act as determinants. When people do not have any vaccination and the majority of people in the province area are not vaccinated, the reasons for people being vaccinated are mostly the requirements of the work unit or

centralized vaccination. Whereas when the number of vaccination needles reached two needles, those who were not inoculated with the third needle were mainly for physical or no reason to prescribe vaccination, nearly nine of those who were inoculated with the third needle were subjectively willing to go and receive the booster needle. Nearly nine percent of those surveyed had positive attitudes towards herd immunity and universal vaccination. That is, the attitude of people toward vaccination has changed from it, and also the phase change illustrates that after a period of coexistence with the COVID-19 outbreak, different anti-disease campaigns and anti plague policies have indispensable roles for the masses. With vaccination rates in China being as high as 85% in

data published in 2022 (Souhu, 2022), this series of facts has some implications for future use when faced with a new pandemic, or in other countries with low vaccination rates.

3.5 Follow-up Investigation After the Booster Vaccination

Hypothesis 5: Side effects may occur after vaccination, but it will not affect people's normal study, work, and life.

A follow-up survey was also done for this survey for those who had already been vaccinated with the booster dose (table 10).

Table 10. questions and options of question 14, 15, 16, and 17

	Classification	Question	Options	exist if...	
14	multiple choice	What are your side effects after the booster injection?	A	dizziness	Q4 chose D
			B	diarrhea	
			C	gastrointestinal discomfort	
			D	poor appetite	
			E	vomiting	
			F	weakness, fatigue, and drowsiness	
			G	runny nose	
			H	muscle soreness	
			I	I do not feel any discomfort	
15	single choice	When did your side effect start after the booster?	A	within half an hour after vaccination	
			B	half a day after vaccination	
			C	one to three days after vaccination	
			D	three days or more after vaccination	
16	single choice	How long did the side effect last after the booster injection?	A	within half an hour	
			B	within half an hour to one hour	
			C	about half a day	
			D	about one day	
			E	about three days	
			F	one week or more	
			B	continuous	
17	single choice	Are you required to stay for half an hour after the booster injection	A	yes	
			B	no	

A total of 38 individuals completed the SARS-CoV-2 vaccine booster vaccination, more than half of whom did not experience any side effects. Among the 14 people with side effects, 12 felt muscle soreness, 4 felt weak and 2 had a small cold. The

main discomfort is muscle soreness. There are no very serious side effects, such as cold and fever. The main discomfort is muscle soreness. There are no very serious side effects, such as cold and fever. Questions 15 and 16 investigated the start time and

duration of side effects in samples inoculated with booster needles. Half of the 14 people felt unwell half a day after vaccination, and nearly 60% of them had side effects lasting no more than one day, which would not have a great impact on people's study and work. The last question (question 17) is asking about whether the vaccination organization is required people to stay for half an hour after the booster injection, all of 38 people are required to stay for half an hour to see if there is any serious side effect. These results indicated that side effects may occur after vaccination, but it will not affect people's normal study, work, and life.

4. Conclusion

The survey collected the data of COVID-19 vaccination, among the sample of students and education in Guangdong, which consisting of objective data, personal feeling and attitude toward the normal vaccine and the booster.

We originally assumed that the vaccination rate is related to the knowledge and understanding of the population about the vaccine, so we designed the question to study the COVID-19 conventional vaccination rate and COVID-19 booster vaccination rate among participants, and then more questions designed to test the subjects' understanding of vaccine. However, the result of Pearson correlation coefficient and Linear Regression indicated that there is no direct correlation between personal understanding and knowledge of vaccine and vaccination rate. Furthermore, the impacts of personal attitude, in other words, whether people believe vaccine work or how easy they can get access to be vaccinated, to the vaccination rate are also concerned. Negative views towards COVID-19 are so common in the internet, which have certain influences, whereas the statistics showed that a large proportion of people holds positive views, which means most of the students and teachers in Guangdong believe that vaccine does work to protect them. It can be inferred that intention of being vaccinated is largely orientated by the government anti-pandemic policies, motivating by the convenient conditions the government was created, such as

increase vaccination points or provides free dozens, which increase the availability to get access to the vaccine.

In addition, the follow-up question of side effects that cause by the vaccination of COVID-19 booster is proved to be normal to most of the people, which it have little influence to border one's study or work.

We can attribute such data to the complexity of vaccine research and development, which requires a variety of technical support, including compliant clinical trials, and various procedures from clinical to market (Nagy & Alhatlani, 2021). The details of vaccine development are not what the public needs and wants to know. For them, what they are most concerned about is what kind of danger they are exposed to and how they need to be protected. Vaccine is one of the protection methods (Shekhar, Garg, Pal, Kottewar, & Sheikh, 2021). People will consider whether this method is effective and how much time and money it will take to obtain it. At this time, the promotion of policy is more important, because it is the key to convey effective information, and also an important link in determining the availability of vaccines. There is a reasonable inference that there are policy influences that are quietly driving citizens in this country closer to the decision about the vaccination.

Reference

COVID-19 Booster. (n.d.). Retrieved February 22, 2022, from baike.baidu: <https://baike.baidu.com>

Dadar, M., et al. (2018). Advances in Designing and Developing Vaccines, Drugs and Therapeutic Approaches to Counter Human Papilloma Virus. *Frontiers in immunology*, 9, 2478. <https://doi.org/10.3389/fimmu.2018.02478>

Kamar, N., et al. (2021). Three Doses of an mRNA Covid-19 Vaccine in Solid-Organ Transplant Recipients. *The New England journal of medicine*, 385(7), pp. 661-662. <https://doi.org/10.1056/NEJMc2108861>

Nagy, A., & Alhatlani, B. (2021). An overview of current COVID-19 vaccine platforms. *Computational and structural biotechnology journal*, 19, pp. 2508-2517. <https://doi.org/10.1016/j.csbj.2021.04.061>

New crown vaccine third needle, officially opened! (2021, 9 23). Retrieved 2 22, 2022, from Souhu: https://www.sohu.com/a/491479920_121124540

Shekhar, R., et al. (2021). COVID-19 Vaccine Booster: To Boost or Not to Boost. *Infectious disease reports*, 13(4), pp. 924-929. doi:10.3390/idr13040084

Sina. (2022, 12, 22). Retrieved February 7, 2023, from news.sina.cn: https://news.sina.cn/zt_d/yiqing0121?luicode=10000011&lfid=1076031768428944

Slifka, K. M., & Amanna, I. (2014). How advances in immunology provide insight into improving vaccine efficacy. *Vaccine*, 32(25), pp. 2948-2957. <https://doi.org/10.1016/j.kint.2021.04.010>

Souhu. (2022, 1 8). Retrieved February 24, 2022, from souhu.com: https://www.sohu.com/a/515207579_161795?g=0

SPSS Online. (2022). Retrieved 2 23, 2022, from SPSS: <http://www.spssau.com>

The SPSSAU project, 21.0. (2021). Retrieved February 26, 2022, from SPSSAU: <https://www.spssau.com>

Vaccine. (2022, 2 10). Retrieved from baike.baidu: <http://baike.baidu.com>

Wang, Y., Xing, M., & Zhou, D. (2020). Coronavirus disease-19 vaccine development utilizing promising technology. *Current opinion in HIV and AIDS*, 15(6), pp. 351-358. <https://doi.org/10.1097/COH.0000000000000648>

The Effects of Different Materials Used in Non-Aqueous Lithium Air Battery Electrolytes on Ionic Conductivity and Durability

Aditya Vishwa¹ *

¹Westview High School, Portland, OR USA

Received October 24, 2022; Revised March 1, 2023; Accepted, March 20, 2023

Abstract

Batteries are increasingly important in moving away from fossil fuels towards a more sustainable future; they are currently used in many different applications, such as cars, portable electronics, and drones, but certain applications of batteries, such as in large aircraft, are limited by their gravimetric energy density. While many high energy density batteries have been synthesized, most lack high durability and are unable to withstand many charge and discharge cycles without losing considerable energy storage capacity. For example, lithium air batteries have one of the highest gravimetric energy densities of any battery, but because lithium ions are not able to flow freely from the anode to the cathode and vice versa. This paper examines recent advancements in lithium air battery materials and explores any potential developments for lithium air battery viability. It also discusses the various issues that plague lithium air battery advancement, and its place in the future of green energy and a viable source of energy storage.

Keywords: Batteries, Energy density, Lithium-air battery, Durability

1. Introduction

Batteries are used in a variety of applications, from small portable electronic devices to EV's. This has partly been due to numerous advances in the volumetric energy densities of batteries, allowing ample amounts of energy to be efficiently stored in a small form factor. But one area where batteries have struggled to be utilized is where light weight high energy storage capabilities are required, such as in the aerospace industry. This is currently an area where battery technology is lacking, as though there have been considerable improvement in volumetric energy densities over the years, gravimetric energy density improvements have been lacking, as there has not been a pressing need for such capabilities, until now, when the world is transition into to a future of renewable energy sources.

Attempts to improve gravimetric energy densities in standard lithium ion batteries have been successful to an extent in that their weight has not hampered their use on a broad scale, but successful attempts to drastically improve gravimetric energy densities to that point where they can compete with gasoline and jet fuels have been limited due to the physical properties of electrode materials as well as the electrolytes. Lithium air batteries, however, relying on a porous oxygen based cathode, are able to yield gravimetric energy densities up to 39.6 MJ/kg, close to that of gasoline (*Lithium-Air Battery - an Overview* | *ScienceDirect Topics*, n.d.). However, as promising as it sounds, lithium air batteries are plagued by numerous issues, the most significant of which is their low cyclability, caused by poor ionic conductivity of the electrolytes used in them.

This paper discusses the basics of lithium ion

* Corresponding Author
Aditya.vishwa05@gmail.com

Advisor: Jaclyn Schillinger
mentors@polygence.org

batteries as well as the advantages and issues related to lithium air batteries and their real world implementation.

2. Findings

2.1 Basic Battery Electrochemistry

To generate current which can be used to power devices, a steady flow of electrons must be generated, which move through such devices, such as a lightbulb. This flow of electrons comes in the form of alternating current, where direction of flow alternates, and direct current, where the flow is in a singular direction, the latter of which batteries provide. In order to generate this flow of electrons, redox reactions must take place within the battery. Redox reactions can be divided into two half-reactions, the oxidation and reduction reactions. These reactions help to convert the chemical energy stored in batteries into electrical energy. In the oxidation reaction, atoms lose electrons and ions. In the reduction reaction, the electrons, and ions, come together, leading to an atom gaining electrons; this process uses up electrons. Essentially the oxidation reactions produce electrons, which are then used up by the reduction reactions.

Due to the opposite nature of these half reactions, they have to occur at separate places within a battery. Batteries typically have four major components; an anode, a cathode, electrolyte solution, and a casing (to protect the battery from external damage). The oxidation half reactions occur at the anode, and the reduction reactions occur at the cathode, the two of which are separated by an electrolyte solution. When the oxidation reaction occurs at the anode, the electrons produced essentially pile up and since electrons are negatively charged, they repel each other and want to move away to a less negative area which would be at the cathode where these electrons can be used up (mischa, 2016). But since the anode and cathode are separated within the battery and electrons are prevented from flowing from the anode to the cathode within the battery, an alternative route can be created where electrons can flow from the anode to the cathode and through this route is where current is created (mischa, 2016). In this route things

that require power can be placed, such as a lightbulb, which the electrons will flow through on their way to the cathode (Path of an Electron through an Electric Circuit, n.d.). But if this route from the anode to the cathode is blocked, nothing occurs as no electrons are moving and no current is generated (mischa, 2016). However, this process is limited to the material available to be used in the reduction reactions within the battery; once those are used up, the redox reactions can no longer occur and the battery is essentially dead; this is what happens in non-rechargeable batteries when they are used (Every Battery Eventually Dies – Here’s Why, n.d.).

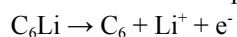
2.2 Conventional Electrolytes

An electrolyte paste prevents the electrons from the anode reaching the cathode where they would be used up (mischa, 2016). The casing just holds the components in place and protects the battery from damage. An electrolyte is a substance that is found in all batteries. They can come in many different forms such as liquids, gels, as well as solids, in some solid-state batteries. An electrolyte is a very important part of the battery because it is what allows the ions to move from the anode to the cathode and vice versa. This is important for a battery's ability to be charged and recharged as the ions from the redox reaction need to be able to get back to their anode and cathode (Components of Cells and Batteries, n.d.). Charging basically reverses the ions back to their starting point, their anode and cathode, however, not all ions make it back and over time this causes the energy capacity to degrade over time, causing it to eventually become unusable (Components of Cells and Batteries, n.d.). In batteries such as lithium air batteries, the issue is with non-aqueous electrolytes, which are very resistant to allowing the products of the redox reactions through, making lithium air batteries almost unusable.

2.3 Lithium Ion Batteries

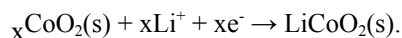
Lithium-ion batteries are relatively high energy density, and most can be recharged. The most common type of lithium-ion battery is a lithium cobalt oxide battery (A Guide to the 6 Main Types of

Lithium Batteries, 2021). In this battery, Li^+ ions, lithium in a +1 oxidation state, move from the anode to the cathode and vice versa through the electrolyte solution. The anode in lithium cobalt oxide batteries contains graphite which helps the batteries have a higher energy density, faster charging, higher durability, and a longer lifespan. The graphite used as the anode is lithiated, meaning the graphite is made of intercalated layers of lithium and graphite with a chemical formula of Li_xC_6 . The cathode in lithium cobalt batteries is made of a lithiated metal oxide, specifically Li_xCoO_2 . However, lithium-ion batteries can also have cathodes made of other materials, such as polyion type cathodes: LiMPO_4 , $\text{LiM}(\text{PO}_4)_F$, $\text{Li}_2\text{MP}_2\text{O}_7$, Li_2MSiO_4 , LiMSO_4F , LiMBO_3 . Lithium-ion batteries can also have cathodes made of spinel oxides such as LiMn_2O_4 . The electrolyte solution should allow the Li^+ ions to move freely between the anode and the cathode. Most lithium-ion batteries today use a liquid or gel electrolyte such as LiPF_6 . Solid-state electrolytes are not currently used commercially because many of the solid-state electrolytes do not allow the Li^+ ions to pass through freely (Bartholome et al., n.d.). During the battery's discharge, the lithium in the lithiated graphite anode is oxidized into Li^+ , and an electron is freed. The chemical reaction happening is



The Li^+ ions move from the anode to the cathode through the electrolyte solution, which will get used up in the reduction reaction at the cathode. The electrons produced during oxidation want to move from the anode to the cathode. However, since electrons take the path of least resistance, they do not move through the electrolyte solution to the cathode. There is also a separating layer in many batteries which prevents the flow of electrons from the anode to the cathode and only allows the Li^+ ions to pass through. In the cathode, where the reduction reaction occurs, the cobalt is reduced to a lower oxidation state, and the Li^+ ion combines with the electrons from the oxidation reaction to form lithium cobalt oxide. Since the electron is required for the reduction reaction to occur, the reduction reaction will not occur until a path for the electrons is given from the anode to the cathode. So, no current will be generated until the battery is plugged in with its positive and

negative terminals connected. The chemical reaction occurring at the cathode is



When the battery recharges, a current put into the battery causes the Li^+ ions to go from the LiCoO back to the lithiated graphite anode (Lithium-Air Battery - an Overview | ScienceDirect Topics, n.d.).

2.4 Solid State Electrolytes

Solid-state batteries are another very high energy-density battery type being researched. They work similar to conventional batteries, except for electrolytes. Conventional batteries use a liquid or gel electrolyte, but these have many issues, such as being more prone to catching on fire and causing them to be heavier (Balshaw2019-11-11T09:38:00+00:00, n.d.). Solid-state batteries can solve these issues with faster charging times, higher durability, and much safer (Mauger et al., 2019). However, solid-state batteries are still being researched, as creating a suitable electrolyte is still causing issues. Solid-state electrolytes, unlike liquid or gel electrolytes, make it much more difficult for ions to pass from the anode to the cathode (Uddin & Cho, 2018).

2.5 Effects of Temperature on Batteries

Materials with a higher electrochemical potential lead to higher energy-density batteries (Liu et al., 2016). Lithium is used in many batteries because of its high electrochemical potential and its lightweight (Goonan, 2012). As the temperature increases, batteries' energy storage capacity increases. However, as the temperature goes down, the energy storage capacity of batteries goes down (BU-502: Discharging at High and Low Temperatures, 2010). So, the temperature is one of the main effects of battery performance: including discharge rate, recharge rate, and energy density (Temperature Effects on Batteries, n.d.). Another major trend is that charging batteries faster can also degrade the batteries faster. Slower charging can increase the lifespan of a battery, and the number of charge cycles can also affect the battery's energy capacity. More charges lead to less energy capacity over time due to

battery degradation (src="https://www.rd.com/wp-content/uploads/2020/07/Brooke-Nelson.jpg?fit=50 et al., 2019).

2.6 Lithium-Air Batteries

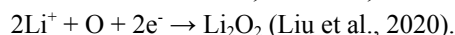
One type of battery with a much higher energy density than lithium-ion batteries is lithium-air batteries. Lithium-air batteries have one of the highest theoretical energy densities of any battery, about five times more than lithium-ion batteries and close to that of gasoline at 43.2MJ/Kg (Liu et al., 2020). Two types of rechargeable lithium-air batteries are being worked on, aqueous and non-aqueous. Lithium-air batteries have the same components as a normal battery: an anode and a cathode, and they rely on reduction and oxidation reactions to generate current. The anode used is pure Lithium, and the cathode consists of carbon and oxygen gas (Liu et al., 2020). The electrolytes used are non-aqueous organic electrolytes, allowing the lithium ions to pass through freely from the anode to the cathode and vice versa. Lithium-air batteries are separated based on their electrolytes: non-aqueous aprotic electrolyte, aqueous electrolyte, solid-state electrolyte, and hybrid electrolyte (Liu et al., 2020). As the battery discharges, the Li^+ ions move from the Lithium to the carbon-oxygen anode by dissolving and moving through the electrolyte. The electrons move to the cathode from the anode through an external path. This process is reversed when the battery is charged as in a conventional battery. The chemical reactions for discharging are



The anode half reaction, oxidation, is



and the cathode half reaction, reduction, is



2.7 Lithium-Air Battery Issues

The issue with this type of lithium-air cell is that the product of the reaction does not dissolve in the electrolyte and ends up depositing in the cathode and, over time, reduces the capacity and durability of the battery due to the usable material within the battery, reducing (Liu et al., 2020). An aqueous electrolyte

can solve this problem by using a lithium-ion conductive water-stable solid electrolyte in which the discharge product is soluble within the electrolyte. However, this results in a lower energy density. The redox reaction for this type of cell is



The air used in these batteries can also affect performance. Moisture in the air can corrode the lithium within the battery.

When operating lithium-air batteries, the ideal conditions would be to operate them using pure oxygen (Asadi et al., 2018). But to make them viable for use on a wide scale in many different industries, they need to be able to work in an environment where they will be exposed to different environmental factors such as heat, cold, humidity, and other impurities in the air.

Humidity for lithium-air batteries can be both beneficial and harmful as well. This can be seen in a study where under the conditions of pure oxygen and zero humidity, the discharge voltage was 2.51 volts. However, under the conditions of 84% humidity, the discharge voltage increased to 2.79 volts. The energy efficiencies also increased from 66.2% to 73.8% from the dry air to the air with 84% relative humidity, respectively (Tan et al., 2016). But this can also be harmful as this can cause a buildup of hydrogen within the battery (Byrne, 2021).

The main issue right now with lithium-air batteries is that when the redox reactions occur within lithium-air batteries, the product, lithium peroxide, is not able to decompose back into lithium oxide (Tan et al., 2016), which reduces battery energy storage capacity (Song et al., 2020).

The product of the redox reactions for lithium-air batteries under high relative humidity is lithium-hydroxide, which requires less energy to decompose than the lithium peroxide created from the dry air. This allows for better batteries because lithium hydroxide allows for a higher discharge voltage when formed (Gallant, 2020).

But other problems can arise from using lithium-air batteries under high humidity environments. This is because pure lithium is highly sensitive to moisture. If excessive moisture comes into contact with the lithium, lithium hydroxide will also be formed, but also hydrogen, which makes the

batteries unsafe due to it being very flammable (Byrne, 2021).

Though it may be possible to create lithium-air batteries with very high gravimetric energy densities, such as up to 11000 Wh/Kg, it may not be practical to be able to use these lithium-air batteries due to other factors that would also have to be considered other than just gravimetric energy density (Lithium-Air Battery - an Overview | ScienceDirect Topics, n.d.). Many lab tests show that lithium-air batteries can reach this gravimetric energy density. However, they have also shown that the use of non-aqueous electrolytes to be able to attain these results has led to speedy degradation in lithium-air batteries, such as only a four charge and recharge cycles, making their gravimetric energy density something that is not very useful as it cannot be sustained (Edge et al., 2021). So, a lower gravimetric energy density battery that lasts longer may be better in the real world due to lithium-air batteries' impracticalities.

But there have been some advancements that could help reduce the degradation of lithium-air batteries. One of the many factors responsible for the degradation of lithium-air batteries is lithium superoxide (LiO_2). Therefore encapsulating, or trapping, the lithium superoxide could increase the recyclability of lithium-air batteries and prevent the degradation of these batteries (Encapsulation as a Method for Preventing Degradation in Li-Air Batteries, n.d.).

But one of the other issues is the practicality of attaining the air needed for them. Many of the impurities in the air used for these batteries can also cause many problems with their durability. The air in lithium-air batteries can have trace amounts of moisture, carbon dioxide, nitrogen, etc. These impurities can cause many side reactions at the anode and the cathode in the battery, causing instability (Liu et al., 2017).

Making lithium-air batteries viable for use with ambient air is an advancement that can make lithium-air batteries viable if the durability issues are fixed. One of the components of ambient air is carbon dioxide. But it can react with oxygen, as well as with lithium, forming lithium carbonate. The advantage of lithium carbonate is that it can provide a way to make lithium-air batteries cyclable, being able to withstand

multiple charger cycles due to the lithium carbonate being easier to decompose. Lithium carbonate, on the other hand, can also be detrimental to lithium-air cells as it can also cause the cells to dry out (Liu et al., 2020).

Lithium-air batteries can last many charge cycles due to recent development, even up to 1200 charger cycles (Kondori et al.). However, the above tests were run at a very low voltage, allowing the lithium-air batteries to last longer than they usually would have. But discharge voltage is vital in battery use as though a battery may be able to hold a lot of power if it cannot discharge it at a reasonable amount, then it is not of any actual use.

3. Discussion

Lithium-air batteries, as any other liquid or gel electrolyte battery, come with risks. The most major of these risks is the risk of the batteries catching on fire due to overheating or leaking, causing the release of many flammable compounds. Lithium-air batteries currently work best under very high humidity conditions, provided water can be kept from damaging certain electronics that manage the battery. The water in the air can react with lithium-ion, forming lithium hydroxide. This helps lithium-air batteries run better because lithium hydroxide is more easily decomposable within the battery, allowing the battery to last longer due to easier charge and recharge cycles being able to run. However, when lithium reacts with water, it forms both lithium hydroxide and hydrogen gas, the latter of which can be very flammable, causing many safety concerns with lithium-air batteries. So, the conditions in which these batteries function the best also yields the most dangerous conditions these batteries are used in. This must be worked out in the future by altering the material used for the electrodes, allowing safer products to form to create safer batteries.

Lithium-air batteries may be viable for use in the future. However, only if substantial research is dedicated to increasing their durability, discharge rates, and energy capacity. Lithium-air batteries cannot last for numerous charge cycles, which is very important for them to become viable. Lithium-ion batteries used in devices that require very little

power, such as TVs or small appliances, can easily be replaced at a low cost. However, much larger lithium-based batteries, such as the ones used in EVs today, are much more expensive to manufacture and replace, with many costing about 15% of the price of the car itself. So, the even more expensive lithium-air batteries must demonstrate a long lifespan in order for them to be economically viable. Though they have many performance benefits, they only provide a small quantity of use other than in aerospace applications and other applications where the weight of the battery is very critical. However, for EVs, battery weight is not as much of a detriment to their range as it would be in planes, and therefore is less of a concern. Due to spatial limitations, volumetric energy density is more critical in many uses today. So, if lithium-air batteries became commercially viable, they would not be used on a vast scale, causing their costs to be much greater than traditional lithium-ion batteries.

The issues with solid-state batteries may be solved before those of lithium-air batteries. Even if lithium-air batteries' durability is improved to viability, they will still rely on liquid or gel electrolytes, which are still not as safe as solid-state electrolytes. A lithium-ion solid-state battery could have a very high energy density, not as high as lithium-air batteries, but still, a high enough gravimetric energy density for them to be revolutionary. A pairing of lithium-air batteries that use solid-state electrolytes would result in the safest and highest energy-density battery. But the cost would preclude this kind of battery as it would only be economically viable in applications of high performance or in aerospace applications where saving weight is a requirement and can result in more range, which would be more cost-effective. There are also many advancements that can result from creating a lithium-air solid-state battery, such as electric planes, which still need to be more viable due to the gravimetric energy density of current batteries being very low when compared with that of jet fuel.

4. Conclusion

The current advancement in lithium-air batteries has slightly increased their durability. There has also

been much research done in all the major areas where problems are in lithium-air batteries: with the electrolytes, the redox products, and the side reactions occurring in lithium-air batteries. But solid-state batteries are currently beating lithium-air batteries to the market and do not require as much development as lithium-air batteries. Moreover, for the large-scale use of lithium-air batteries, especially in aerospace operations, these batteries would have to have a very high discharge rate and be able to sustain how much energy they can hold over long periods. Since lithium-air batteries are still struggling first to become usable, their discharge rate has yet to be explored that much, and more research would have to be done to make them more commercially viable in the future.

There is currently some successful research into making lithium-air batteries last for multiple charge cycles, making them viable for commercial use (Kondori et al.). Specifically, getting them to last for 1200 charger cycles, in some cases, is very close to how many charge cycles current lithium-ion batteries, used in, for example, electric cars, last. However, this testing was done with meager discharge rates, causing the number of charger cycles the batteries lasted to be much higher than in practical use. The discharge rate is critical because it is about how many enemies a battery can hold and how much it can release at a time. If there were an electric car with a range of 600 miles but a max speed of 5mi/hr., then it would not really be of much use. This is still the current dilemma lithium-air batteries face. However, it is progress that these batteries can at least last more than a couple of charger cycles without dying.

References

- A Guide To The 6 Main Types Of Lithium Batteries.* (2021, September 27). Dragonfly Energy. [https://dragonflyenergy.com/types-of-lithium-batteries-guide/#:~:text=Lithium%20cobalt%20oxide%20\(LCO\)%20batteries%20are%20used%20in%20cell%20phones](https://dragonflyenergy.com/types-of-lithium-batteries-guide/#:~:text=Lithium%20cobalt%20oxide%20(LCO)%20batteries%20are%20used%20in%20cell%20phones)
- Asadi, M., et al., (2018). A lithium–oxygen battery with a long cycle life in an air-like atmosphere. *Nature*, 555(7697), 502–506. <https://doi.org/10.1038/nature25984>

Balshaw L. (2019) Gel polymer electrolyte cuts risk of battery fires. *Chemistry World*. Retrieved 2022, from <https://www.chemistryworld.com/news/gel-polymer-electrolyte-cuts-risk-of-battery-fires/4010625.article>

Bartholome, T., Hankins, K., & Keller, N. (n.d.). *Lithium Ion Batteries What are lithium ion batteries and how do they work?* <https://doi.org/10.1063/1.3047681>

Batteries, circuits, and transformers - U.S. Energy Information Administration (EIA). (n.d.). [www.eia.gov](https://www.eia.gov/energyexplained/electricity/batteries-circuits-and-transformers.php#:~:text=Batteries%20produce%20electricity). Retrieved 2022, from <https://www.eia.gov/energyexplained/electricity/batteries-circuits-and-transformers.php#:~:text=Batteries%20produce%20electricity>

Bertrand, G. (2020). *About Batteries*. Mst.edu. <https://web.mst.edu/~gbert/BATTERY/battery.html#:~:text=The%20battery%20operates%20through%20electrochemical>

BU-502: Discharging at High and Low Temperatures. (2010, September 1). Battery University. <https://batteryuniversity.com/article/bu-502-discharging-at-high-and-low-temperatures>

Byrne, A. (2021, January 7). *H2 Hydrogen Detection in Battery Rooms*. Eagle Eye Power Solutions. <https://eepowersolutions.com/hydrogen-detection-and-control/>

Components of Cells and Batteries. (n.d.). Depts.washington.edu. Retrieved 2022, from <https://depts.washington.edu/matseed/batteries/MSE/components.html#:~:text=The%20Electrolyte%20is%20the%20>

Edge, J. S., et al., (2021). Lithium ion battery degradation: what you need to know. *Physical Chemistry Chemical Physics*, 23(14), 8200–8221. <https://doi.org/10.1039/d1cp00359c>

Encapsulation as a method for preventing degradation in Li-air batteries. (n.d.). MIT News | Massachusetts Institute of Technology. <https://news.mit.edu/2022/encapsulation-method-preventing-degradation-li-air-batteries-0120>

Every Battery Eventually Dies – Here's Why. (n.d.). Pale Blue. Retrieved 2022, from <https://paleblueearth.com/blogs/news/every-battery-eventually-dies-heres-why>

Gallant, B. M. (2020). Unlocking Reversibility of LiOH-Based Li-Air Batteries. *Joule*, 4(11), 2254–2256. <https://doi.org/10.1016/j.joule.2020.10.018>

Goonan, T. G. (2012). *Lithium Use in Batteries*. https://pubs.usgs.gov/circ/1371/pdf/circ1371_508.pdf

HowStuffWorks Authors. (2000, September 11). HowStuffWorks. <https://www.howstuffworks.com/about-author.htm>

Ling, J., et al., (2021). Phosphate Polyanion Materials as High-Voltage Lithium-Ion Battery Cathode: A Review. *Energy & Fuels*, 35(13), 10428–10450. <https://doi.org/10.1021/acs.energyfuels.1c01102>

Lithium-Air Battery - an overview | ScienceDirect Topics. (n.d.). [www.sciencedirect.com](https://www.sciencedirect.com/topics/engineering/lithium-air-battery#:~:text=The%20). Retrieved 2022, from <https://www.sciencedirect.com/topics/engineering/lithium-air-battery#:~:text=The%20>

Liu, C., Neale, Z. G., & Cao, G. (2016). Understanding electrochemical potentials of cathode materials in rechargeable batteries. *Materials Today*, 19(2), 109–123. <https://doi.org/10.1016/j.mattod.2015.10.009>

Liu, T., et al., (2020). Current Challenges and Routes Forward for Nonaqueous Lithium–Air Batteries. *Chemical Reviews*, 120(14), 6558–6625. <https://doi.org/10.1021/acs.chemrev.9b00545>

Liu, Y., et al., (2017). Understanding and suppressing side reactions in Li–air batteries. *Materials Chemistry Frontiers*, 1(12), 2495–2510. <https://doi.org/10.1039/c7qm00353f>

Mauger, A., et al., (2019). Building Better Batteries in the Solid State: A Review. *Materials*, 12(23), 3892. <https://doi.org/10.3390/ma12233892>

Bhatt, A., et al., (2016, February 25). *How a battery works*. <https://www.science.org.au/curious/technology-future>

/batteries#:~:text=The%20electrolyte%20provides%20a%20medium

Path of an electron through an electric circuit. (n.d.). Physics Stack Exchange. Retrieved 2022, from <https://physics.stackexchange.com/questions/57611/path-of-an-electron-through-an-electric-circuit>

Pentland, W. (n.d.). *Lithium-Air Battery Breakthrough May Mean Game Over For Gasoline.* Forbes. <https://www.forbes.com/sites/williampentland/2015/10/31/lithium-air-battery-breakthrough-may-mean-game-over-for-gasoline/?sh=1d306a916638>

Song, L.-N., et al., (2020). Tuning lithium-peroxide formation and decomposition routes with single-atom catalysts for lithium–oxygen batteries. *Nature Communications*, 11(1). [https://doi.org/10.1038/s41467-020-15712-zsrc="http://www.rd.com/wp-content/uploads/2020/07/Brooke-Nelson.jpg?fit=50, img class="avatar" alt="Brooke N., Feb. 04, 50" width="50" height="50"> Brooke N., & 2022. \(2019, September 26\).](https://doi.org/10.1038/s41467-020-15712-zsrc=)

The Best Way to Charge Your Device Will Make Its Battery Last Way Longer. Reader's Digest. <https://www.rd.com/article/charge-smartphone-laptop-battery-last-longer/>

Tan, P., et al., (2016). Effects of moist air on the

cycling performance of non-aqueous lithium-air batteries. *Applied Energy*, 182, 569–575. <https://doi.org/10.1016/j.apenergy.2016.08.113>

Temperature effects on batteries. (n.d.). Intercel Services B.V. <https://www.intercel.eu/frequently-asked-questions/temperature-effects-on-batteries#:~:text=Even%20though%20battery%20capacity%20at>

Uddin, M.-J., & Cho, S.-J. (2018). Reassessing the bulk ionic conductivity of solid-state electrolytes. *Sustainable Energy & Fuels*, 2(7), 1458–1462. <https://doi.org/10.1039/c8se00139a>

We won't have electric airplanes until battery tech improves. (n.d.). Engadget. Retrieved July 20, 2022, from <https://www.engadget.com/we-wont-have-electric-airplanes-until-battery-tech-improves-160058357.html>

What is a Battery? (n.d.). Depts.washington.edu. <https://depts.washington.edu/matseed/batteries/MSE/battery.html>

What is self-discharge? (2018, February 20). Panasonic Batteries. <https://www.panasonic-batteries.com/en/faq/what-self-discharge#:~:text=A%20minuscule%20amount%20of%20the>

LED Array Development: Highlights and Challenges of Red Light Therapy

Hilary Pou Yiu Yeung^{1*}

¹St. Clement's School, Toronto, ON CANADA

Received August 30, 2022; Revised March 21, 2023; Accepted, April 5, 2023

Abstract

Red light therapy (RLT) is a recent development that faces much uncertainty due to its novel nature. In essence, it has been found that the light photons trigger respiratory chain components which in turn activate a signaling cascade, and thus result in a cellular response. Some of these biological reactions aid in healing, skin repair, inflammation reduction, and functional improvement. Although such knowledge has been retained, scientists still struggle to understand how RLT truly works and the risks that are paired along. Though limited research has been conducted, this form of therapy allows for a wide variety of applications due to the numerous dosimetry parameters. As parameters have countless combinations, RLT can be used for people of varying ages, skin conditions, illnesses, and health backgrounds. Though seeming positive, such wide combinations leave room for the unpredictability of specific settings that pertain to each user. An additional concern is a basic accessibility; RLT has yet to be integrated into mainstream health improvement. In all, RLT provides optimistic solutions for high-performing humans, though much more research is required to uncover all the risks.

Keywords: Red light therapy, Photons, Skin treatment, Therapy, Healing, Inflammation, Health

1. Introduction

Recent scientific developments like Red Light Therapy (RLT) have increased the general well-being of high-performing people such as athletes and trainers. This form of physical therapy has been discovered to “treat a multitude of conditions that require stimulation of healing, relief of pain and inflammation, and restoration of function” (Avici et al., 2013). Despite these many benefits, however, controversy flows around the topic due to the uncertainty “about fundamental molecular and cellular mechanisms responsible for transducing signals from the photons” (Avici et al. 2013) as well as the “significant variations in terms of dosimetry

parameters” (Avici et al., 2013) needed for phototherapy. Since this type of phototherapy has yet to be accepted into the mainstream world of, “science and medicine” (Sontea et al., 2013), there is less access. Although these issues prevail, the idea of red light therapy is simple. Pagán (2021) from WebMD simply explains it as how the mitochondria, “soaks up [energy] from red light”.

The history of RLT dates back to the late 19th century by a Danish physician named Niels Finsen, who developed RLT in 1896 as a treatment for lupus vulgaris, a type of tuberculosis (Vorwaller, 2020). Vorwaller (2020) reports Dr. Finsen’s historical process as using electric light a total of 15 times, each time focused on the same 2cm of skin. Finding

* Corresponding Author
ms.hilary.yeung@gmail.com

Advisor: Jessica Sands
jls642@cornell.edu

immense success, Dr. Finsen was awarded the Nobel Prize of Medicine following this utilization (Lightwave, n.d.) in 1903. Light therapy research ramped up in the 1960s in Eastern Europe to heal “chronic pain, arthritis and associated conditions, joint rehabilitation, and soft-tissue injury along with other medical ailments” (Lightwave, n.d.). After word spread about the success and effectiveness of RLT, it is not surprising that this method became a popular skin treatment “given the claims that it did not damage your skin surface and had many benefits that promoted skin health,” (Team Nutrisense, 2022).

While the pandemic has unfolded prevailing health issues for citizens of the world, considering new technological development like RLT is essential in aiding the evolving human body. For instance, there have been claims that this form of therapy will aid in curing severe illnesses such as the Niemann-Pick disease purely due to the properties of light (Schwarcz 2022). RLT is a convenient, effective, low-cost, and low-risk solution for populations such as high-performing athletes and seniors. For instance, seniors may experience skin problems, typically brought with old age; RLT can provide a simple solution to energy rejuvenation and wound repair. All in all, “RLT is generally safe and may be a very effective treatment option for people seeking smaller changes in their skin or to keep the skin healthy and reduce inflammation” (Medical News Today, 2019). This paper will explore more about RLT and the effect of LED array development on the advantages and risks. As with any development, scientists currently battle out the pros and cons of RLT and its applications in hopes to build technology that advances humanity.

2. Internal Action

While the results of RLT can be significant, it is critical to understand how this form of phototherapy works. Victor Sontea (2013) and his colleagues explored the mechanism of RLT in their 2013 E-Health and Bioengineering Conference. Sontea et al. (2013) explain how the photo-bio-modulation, a form of light therapy that utilizes non-ionizing light sources (Light Force Therapy Lasers), by red light, “has been ascribed to the activation of mitochondrial

respiratory chain components resulting in [the] initiation of a signaling cascade that promotes cellular proliferation and cytoprotection” (Sontea et al., 2013). In short, this means that the photons from the red light trigger a cellular response in the cells, leading to the advantages of RLT. On the contrary, discussions around the uncertainty of how the cascading signals actually work have been a topic of controversy. Presently, signaling cascades are a topic of scientific interest where researchers are working to confirm the directions and effects of each cascade in a definite manner (Catozzi et al., 2016). Nonetheless, studies have shown that “proposed underlying mechanisms include the photostimulation of terminal molecules in the electron transport chain and the subsequent adenosine triphosphate (ATP) concentration increase” (Wunsch et al., 2014). Likewise, low-level laser light therapy, which includes RLT, activates, “electron transport, ATP nitric oxide release, blood flow, reactive oxygen species increase and diverse signaling pathways” (Avici et al., 2013). As a result of higher ATP production, biological systems can work much more efficiently, leading to the numerous advantages of RLT.

Overall, in order for RLT benefits to be present, the light must initiate a chain reaction that prompts other organs within one’s body, resulting in the highlights of RLT that are discussed later on in this paper. By calling on molecules that eventually go through cellular respiration, energy production is generated at a higher level. Notably, the human body carries energy through ATP, which is then used for processes such as muscle contraction, metabolism, growth, and repair (Dunn & Grider, 2022).

3. Conditional Improvement

After comprehending the RLT effects on our body, we can now dive into the physical improvements that this form of phototherapy creates. As LED developments are constantly evolving, benefits such as “lower cost,...absence of heat, and [large] arrays for large wound treatment,” (Sontea et al., 2013) have been found. Although scientists believe that blue light is just as effective as red light, Sontea et al. (2013) argue that “light close to and in

the red and infrared range effectively promotes wound healing”. Likewise, Avici et al. (2013) explain how “red light is known to penetrate deeper in tissues when compared to blue light”. As suggested by both papers, red light seems to be more effective in healing in comparison to blue light. Important to note, however, is that combining both blue and red light can produce an “overall decrease melanin level” (Avici et al., 2013, p.10), which produces depigmentation and can increase one’s risk of sunburns. Nevertheless, the regenerative capabilities of red light on its own have other advantages, including dental pain reduction, hair loss improvement, and tendonitis recovery (Pagán, 2021). With the introduction of LED devices, these treatments are much easier to apply as the devices have, “simplified the application to large areas of skin” (Avici et al., 2013, p.12).

A study by Wunsch & Matuschka (2014) explored the RLT application on 136 volunteers and it was noted that there was a significant experience in improvement of personal assessments of, “skin feeling and complexion in clinical outcomes as assessed by collagen density and skin roughness measurements and in the reduction of fine lines and wrinkles. Wunsch & Matuschka (2014) explain their method by saying the following: “113 subjects randomly assigned into four treatment groups were treated twice a week with either 611–650 or 570–850 nm polychromatic light (normalized to ~9 J/cm² in the range of 611–650 nm) and were compared with controls (n=23)”. Moreover, 30 sessions were implemented for each subject with appropriate evaluation. This study evidently shows that RLT provides a safe, “atraumatic photobiomodulation treatment of skin tissue with high satisfaction rates” (Wunsch & Matuschka, 2014). Clearly, light-emitting diodes (LEDs) are effective light sources that have advantages such as broad beam width, and the “expected signal strength given the direction and radiation distance of an antenna” (A. H. Systems) and cost efficiency (Rohringer et al., 2017).

Overall, such sources emphasize the benefits that come with RLT, both in an efficient and reliable manner. Though researchers clearly believe in this solution, there seems to be a concern surrounding the

temporary-like expectation that RLT creates. As parameters of RLT (further discussed in the following section) create countless combinations, one must wonder how a patient can confirm the appropriate number of RLT sessions, before a burn occurs.

4. Complications and Accessibility

With any new developments, there are also risks, and RLT is no different. With limited time and knowledge of the novel solution, this form of phototherapy comes with numerous complications. Firstly, because RLT has such a wide beamwidth, there can be, “significant variations of...dosimetry parameters [such as] wavelength, irradiance or power density, pulse structure, coherence, polarization, energy, fluence, irradiation time, contact vs non-contact application, and repetition regimen” (Avici, et al., 2013). As a result of the infinite variations, there is a high probability of patients receiving burns (Avici, et al., 2013). Notably, Avici et al. (2013) view this form of phototherapy with criticism, a lens that is respected across the scientific field in order for a reliable application like RLT to be distributed across the world. Access to this type of treatment is also an additional concern of RLT. Despite the numerous pieces of evidence that display the success of RLT, this form of phototherapy has yet to be implemented into the mainstream of medicine which, “makes it unavailable to patients who could benefit from it.” (Sontea et al., 2013). Many studies have been conducted without, a “proper understanding of biological effects of light...[and] scientific methodology” (Sontea et al., 2013) due to the novel nature of RLT. In short, the recent development of LED arrays is surrounded by an insatiable understanding of the true effects of RLT and the lack of time for conducting thoroughly planned experiments. Many scientists have inquired if the combination of blue and red light would significantly improve patient recovery, for instance. For example, Avici et al. (2013) discuss the recent usage of phototherapy for acne treatment and their paper states that “mainly blue light, red light or combination of both” can be used, emphasizing the uncertainty around the appropriate wavelength to be used for specific treatments, such as acne. As light

consists of a wide range of frequencies, often ranging from 400 to 700 nanometers, further research is critical to solidify the appropriate range where humans can use RLT safely.

That being said, scientists argue that the advantages of RLT overpower the risks. Nutrisense, a certified program that analyzes wellness states that “[l]ight therapies like RLT work by causing a controlled impact on your skin to promote tissue repair, rejuvenate your cells, and stimulate regeneration. Because this sort of light therapy doesn’t penetrate the skin more than five millimeters, it’s believed not to cause any long-term damage” (2022). This form of therapy has been used by people all across the world to help with rejuvenation and tissue repair, displaying a promising solution for the future of skin health.

5. Conclusion

Red light therapy is a recent scientific development that uses LED arrays to help with many recovery processes. This form of phototherapy can be used, “to treat a multitude of conditions that require stimulation of healing, relief of pain and inflammation, and restoration of function,” (Avici et al., 2013). These solutions are possible due to an internal cell relationship where the absorption of light initiates a signaling cascade which allows the body to react in a certain way. To list a few, the photons help enhance enzyme activity, mitochondrial respiration, and ATP production (Avici et al., 2013). As all of these cellular responses are triggered by infrared light, phototherapy “effectively promotes wound healing” (Sontea, et al., 2013). In addition to these benefits, RLT provides a safe, variable, and cost-efficient solution for anyone to improve skin complexity, reduce inflammation, and stimulate overall healing. On the contrary, recent developments in LED arrays have risks as well. Firstly, the unknown nature of how phototherapy affects our bodies creates uncertainty and controversy around the proper use of RLT. When distinct answers are not provided, scientists are hesitant to claim that RLT is successful, which in turn projects wariness to patients who are curious to use RLT. Secondly, the wide variety of light parameters can create probable burns

on patients as a particular setting on one person may not work on another. Lastly, access to RLT is an issue because of the unconventional belief surrounding this form of phototherapy. Without proper research, patients are unable to access RLT in a consistent manner. Subsequently, future researchers should consider testing the effects of RLT and its different parameters with willing volunteers to test the effectiveness of this novel solution. Experimenting with different wavelengths, minutes per session, and an overall number of sessions will help catalyze and confirm the science behind RLT. Furthermore, researchers should consider using RLT for high-performing athletes to observe how effective this form of light therapy is in energy reproduction, which would be easily seen in athletes. In conclusion, the recent development of RLT result in thorough scientific conversations, where pros and cons are battled out in hopes to build technology that advances humanity.

References

- A.H. Systems. (n.d.) *The Importance of Antenna Beamwidth in RF Testing*. A. H. Systems.
- Avici, P., et al. (2013). Low-level laser (light) therapy (LLLT) in skin: stimulating, healing, restoring. *Seminars in cutaneous medicine and surgery*, 32(1), 41-52.
- Catozzi, S., et al. (2016). Signaling cascades transmit information downstream and upstream but unlikely simultaneously. *National Center for Biotechnology Information, U.S. National Library of Medicine*. DOI: 10.1186/s12918-016-0303-2
- Dunn, J. & Grider, M. H. (2022). Physiology, adenosine triphosphate. *National Center for Biotechnology Information, U.S. National Library of Medicine*.
- Lightwave. (n.d.). *What is the history of LED therapy?* Lightwave.
- Pagán, C.N. (2021 Novemeber). What is Red Light Therapy? *WebMD*.

Rohringer, S., et al. (2017). The Impact of Wavelengths of LED Light-Therapy on Endothelial Cells. *Nature News*, *Nature Publishing Group*, 7(10700). DOI:10.1038/s41598-017-11061-y

Sontea, V., et al. (2013). Effects of the low level light therapy on skin wound using LED. *2013 E-Health and Bioengineering Conference (EHB)*. DOI: 10.1109/EHB.2013.6707243.

Team Nutrisense. (2022). *What is red light therapy?* Nutrisense.

Vorwaller, J. (2022). *A brief history of red light therapy*. Theralight.

Wunsch, A., & Matuschka, K. (2014). A controlled trial to determine the efficacy of red and near-infrared light treatment in patient satisfaction, reduction of fine lines, wrinkles, skin roughness, and intradermal collagen density increase. *National Center for Biotechnology Information, U.S. National Library of Medicine*.

The Relevance of Semaglutide in Obesity and Diabetes

Jonathan Yoo^{1*}

¹Hawthorne Christian Academy, Hawthorne, NJ USA

Received October 9, 2022; Revised March 28, 2023; Accepted, April 18, 2023

Abstract

In spite of strides made in the field of medicine in relation to diabetes, there still remains a prominent percentage of people in the world who suffer from either Obesity or Type 2 Diabetes. Presently, there is no cure for these diseases, and the need for change is clear; therefore, medicine that can help curb the numbers of both diseases is crucial. Fortunately, Semaglutide is a new breakthrough in weight loss medication and has potential to treat Type 2 Diabetes patients as well. This paper aims to explain in depth the history of weight loss treatments, diabetes treatments, and the mechanisms of insulin and Semaglutide. Furthermore, to put this treatment in context, this paper dives deeply into the subtopics of continuous glucose monitoring, glucose management, current insulin research, and present treatment and preventative measures for diabetes today. Drawing from an array of studies, statistics, and scholarly articles, this paper shows the relevance of Semaglutide in a world with rising Obesity and Type 2 Diabetes, and the possible contributions it could have to these widespread diseases. Further research studies and analysis are being conducted daily in hopes of solidifying these first steps towards a cure and prevention.

Keywords: Diabetes, Glucose monitoring, Insulin research, Treatment

1. Introduction

America is one of the most overweight countries in the world, with 33% of American adults classified as obese (IBIS World, 2022). Many studies have shown the negative impacts of elevated BMI (Body Mass Index), such as an increased risk of mortality and morbidity (Padwal et al., 2011). According to one comprehensive study, “each five-unit increment in BMI above 25 kg/m² is associated with increases of 29% for overall mortality, 41% for vascular mortality and 210% for diabetes-related mortality” (Prospective Studies Collaboration, 2009). There are many complications that come with obesity, one of the many includes type two diabetes. Type two diabetes is caused by obesity because of the increase in fatty acids and inflammation in the body (Boden,

2006). As fatty acids and inflammation go up, insulin resistance also goes up, so the body rejects its own insulin, leading to type two diabetes. More than 37 million Americans have diabetes (about 1 in 10), and approximately 90-95% of them have type 2 (the non-inherited form of) diabetes (CDC, 2022). Diabetes is also comorbid with a number of diseases such as neuropathy, heart diseases, renal failures, and vascular complications (Klimek et al., 2015).

Despite the current "pandemic" of diabetes and obesity in the US, current treatments are not efficient nor long-lasting. This paper aims to compare the pros and cons of past and current treatments for diabetes. Since one of the first steps often recommended to prevent or treat this disease is for an individual to lose weight, methods of weight loss are also explored. Lastly, the new drug semaglutide as a

* Corresponding Author
jonathanyoo2006@gmail.com

Advisor: Heather Killeen
am.HeatherKilleen@gmail.com

potential treatment for Obesity and Type 2 Diabetes is analyzed here.

1.1 Obesity and Weight Loss Treatment

Obesity can have many negative health complications, such as high blood pressure, high cholesterol, coronary heart disease, and most prominently, diabetes. Frequently, people have adopted what could be considered unhealthy habits to control weight. These include, but are not limited to, meal-skipping, fasting, smoking for appetite reduction, intense exercise, as well as consuming stimulants such as caffeine, ephedrine, prescription drugs, and energy drinks (NIDDK, 2022). It has been reported that roughly 12% of adolescent women and 8% of men engage in extreme weight control behaviors including taking diet pills, laxatives, diuretics, or purging (Ferraro et al., 2015).

Many businesses have advertised a magic solution to cure weight problems. Some prominent companies include Sensa, HCG Platinum, Double Shot, and "Get High School Skinny". These companies seem enticing to customers who do not want to exercise and those who are impatient to see their results. For example, a weight loss company called Sensa advertised the loss of 30 pounds in only six months by sprinkling food with sweet and savory crystals that were meant to keep people full (Smith, 2014). Although these promises sound enticing, there is no solid evidence for the validity of this product, nor any specific diet and fitness guidelines for this product, and the few clinical studies available are not published in reliable medical journals. Other companies boast of dermal products that claim to reduce belly fat by just applying the product to the skin. For instance, slimming patches such as those produced by Myntlife and Isumi are growing in popularity but are not backed with any scientific evidence of drastic weight loss (Wright, 2022). These products are not approved by the FDA, and do not have the proper studies to fortify their claims. Lawsuits against companies such as "Get High School Skinny" have led to charges of deceptive advertising (Federal Trade Commission, 2014), but many companies (and marketing teams) continue to thrive.

The FDA has approved five weight loss drugs — orlistat (Xenical, Alli), phentermine-topiramate (Qsymia), naltrexone-bupropion (Contrave), liraglutide (Saxenda), and semaglutide (Wegovy) — for long-term use. A deeper dive into the mechanisms of these drugs can help those interested in pharmaceutical options. All of these drugs fall under the GLP-1 receptor agonist class, which encourages the interaction of GLP-1 and GLP-1R. This reaction activates the cyclic adenosine monophosphate (cAMP) signaling pathway which activates protein kinase A (PKA) and Rap guanine nucleotide exchange factor 4 (RAPGEF4, also known as EPAC2). PKA activation then leads to an increase in the intracellular calcium concentration, thereby promoting the mitochondrial synthesis of ATP, and the release of insulin particles into the blood through exocytosis (when a cell transports molecules out of the cell)(Zhao et al 2021). Outside of their GLP-1 agonist characteristics, they each also have some specific mechanisms that are worth comparing. Orlistat covalently binds to the serine residues of active sites of lipases and inactivates them (Bansal, 2022). The inactivation of lipases prevents the hydrolysis of triglycerides, and therefore free fatty acids are not absorbed (Bansal, 2022). Phentermine-Topiramate is actually an anticonvulsant that lowers the seizure threshold by acting on high-voltage-activated calcium channels and voltage-gated sodium channels and has an augmenting effect on GABA-A receptors, which have a side-effect of suppressing appetite (Johnson, 2022). Naltrexone-bupropion is a chemical combination that is theorized to work synergistically in the hypothalamus and the mesolimbic dopamine circuit to promote satiety, reduce food intake, and enhance energy expenditure (Sherman et al., 2016). Liraglutide increases intracellular cyclic AMP (cAMP) leading to insulin release in the presence of elevated glucose concentrations (WCG CenterWatch, 2022). As a result, it decreases glucagon secretion in a glucose-dependent manner. Unfortunately, Liraglutide has also been associated with an increased risk of side-effects including inflammation of the pancreas, kidney and gallbladder problems (Victoza, 2022). Finally, Semaglutide works similarly to this last drug, in that it reduces blood glucose

through a mechanism where it stimulates insulin secretion and lowers glucagon secretion, both in a glucose-dependent manner (WCG CenterWatch, 2022). It is worth noting, that all agents in the GLP-1 receptor agonist class (Glucagon like peptide-1 receptor), are associated with gastrointestinal adverse reactions such as nausea, abdominal pain, and vomiting (Hughes, S & Neumiller, J.J., 2020). Subcutaneous Semaglutide is most commonly associated with GI side-effects, however, more rare side effects have also been reported which should be noted carefully by any patients taking the drug (Mayo Clinic).

Semaglutide, also known by its brand name as “Wegovy”, is the first approved drug for chronic weight management in adults with general obesity or overweight since 2014 (USFDA, 2021). Semaglutide is an interesting weight loss candidate because while it boasts improvements in weight loss management it has also been studied to help manage symptoms of Type 2 Diabetes. FDA approved in 2021, Semaglutide comes as a solution (liquid) in a prefilled dosing pen to inject subcutaneously (under the skin). Semaglutide is in a class of medications called incretin mimetics. These work by allowing the pancreas to release an adequate amount of insulin depending on the circumstances (Medline Plus, 2022). Importantly, insulin allows sugar to move from blood into other body tissues where it is stored and used for energy. Semaglutide injections also work by slowing the movement of food through the stomach and may decrease appetite to result in weight loss, as well as reducing the amount of insulin needed (Medline Plus, 2022). Remarkably, Semaglutide has been shown to be effective with only weekly injections without regard to meals (USFDA, 202). Further, a study that was published in 2022 concluded that once-weekly subcutaneous Semaglutide compared with once-daily subcutaneous liraglutide (another incretin mimetic), resulted in significantly greater weight loss at 68 weeks (Monaco, 2022). This drug was targeted at those with a body mass index (BMI) of 27 kg/m² or greater who have at least one weight-related ailment or in patients with a BMI of 30 kg/m² or greater. It has been proposed that this drug could also be useful to those with both type one and type two diabetes, by managing insulin and

glucose levels, and is currently undergoing clinical trials (Tsoukas, 2022).

1.2 Diabetes and Diabetes Treatment/Glycemic Control

Diabetes and Glycemic Control

Diabetes is characterized by unhealthy levels of glucose in the blood due to poor insulin production or release. Injectable insulin is the most common solution to high glucose levels or poor glycemic control. Glycemic control refers to one’s blood glucose levels. Blood glucose levels drop when we are hungry, or have gone a long time since our last meal, and increase after we consume food. Under normal conditions, a person’s blood glucose levels are not expected to fluctuate too much since as blood glucose increases, so does our insulin level which helps allow glucose to enter cells and leave the bloodstream (Medical News Today, 2022)(Figure 1).

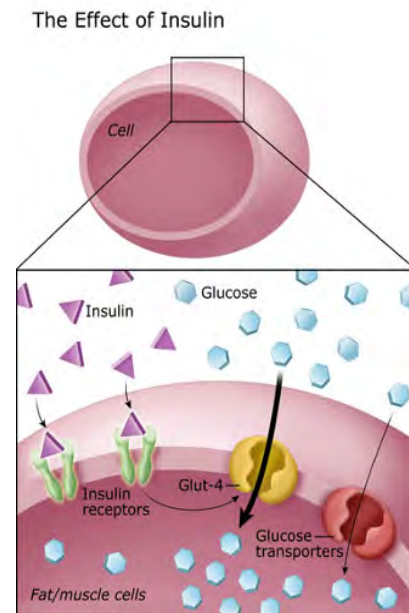


Figure 1. Diagram of insulin and how it reacts with glucose. After food is digested, glucose is released into the bloodstream. In response, the pancreas secretes insulin, which directs the muscle and fat cells to take in glucose. Cells obtain energy from glucose or convert it to fat for long-term storage.(Image source: Diabetes Teaching Center).

Your liver and muscles can take up glucose either for immediate energy or to be stored as glycogen

until it's needed. If the pancreas cannot produce the right amount of insulin to break down the glucose, it will lead to hyperglycemia which is high blood sugar. We also utilize a hormone called Glucagon which is released when our blood sugar is too low and signals to cells to allow sugar (and other fuel) to be released back into the bloodstream (Medical News Today, 2022; Diabetes Teaching Center).

Diabetics deal with two different types of complications of their disease in terms of blood sugar levels: Hypoglycemia and Hyperglycemia, more often known as low and high blood sugar, which can cause serious long term effects if not treated. Hypoglycemia (low blood sugar) has few direct links to critical long-term effects, but they can increase a person's vulnerability to other conditions, such as heart disease, eye disease, kidney disease, and nerve damage (Millar, 2021). Hyperglycemia (high blood sugar) may cause a range of health problems to develop, including skin complications, eye complications, nerve damage, and diabetic ketoacidosis (Felman, 2021). Even though there is no cure for Diabetes, patients can still take measures to treat and prevent their blood sugar from fluctuating.

For most people, blood sugar is an aspect of our diet/biology that is neglected or below our radar, because it is commonly said that if a person does not have diabetes, their sugars will stay level, but this is not always the case. A study in 2018 showed that people without diabetes still had blood sugar spikes (Lily, 2019). In this study, researchers recruited 57 people without diabetes to wear a CGM (Continuous Glucose Monitor) for a "few" weeks where the wearer's blood sugar was monitored every five minutes. The results of this study showed that even non-diabetic people often have blood sugar spikes, some more severe than others, leading to neglected glycemic control. Additionally, the figure below shows an experiment where blood glucose levels were measured after a meal in diabetic and non-diabetic rats (Figure 2). Results showed that after meals, both the diabetic and non-diabetic rats' blood sugars spiked to roughly 300 mg/dl, however non-diabetic and insulin-treated diabetic rats recovered from high blood sugar faster than those without diabetes (Figure 2).

While insulin plays a big role in regulating

healthy glucose levels, it cannot fully prevent blood sugar spikes. Blood sugar spikes happen when the body cannot utilize and absorb the insulin fast enough after the body has ingested food, especially carbohydrates. This little gap in the insulin activation process leads to blood sugar spiking during that short amount of time. Even short blood sugar spikes can have negative effects. In the short term, they can cause lethargy and hunger. Over time, the body may not be able to lower blood sugar effectively, which can lead to type 2 diabetes (Smith, 2022). Frequent blood sugar spikes for diabetics can cause diabetic ketoacidosis (DKA), a potentially deadly condition that causes the blood to become too acidic (Cleveland Clinic).

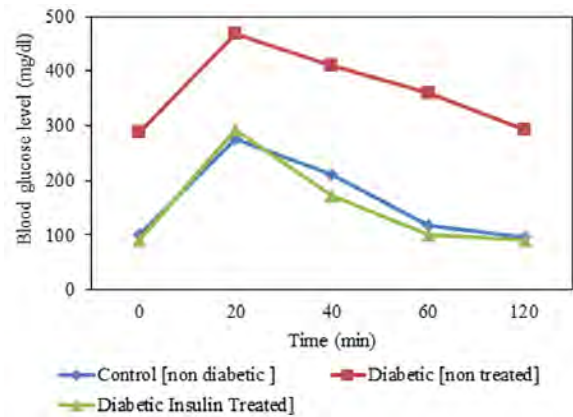


Figure 2. Blood glucose level (mg/dl) using oral Glucose Tolerance Test (OGTT) in Diabetic (red and green) and non-Diabetic (blue) Albino Rats. Time on the Y-axis refers to the time after a meal is given. The Diabetic rats were split into two groups, one treated with Insulin (green) and the other not (red). This graph shows that the diabetic, non-treatment group (red) exhibited much higher blood glucose levels than the other two groups (Figure source: Jato et al., 2018).

Furthermore, consistent blood sugar spikes can lead to insulin resistance, causing longer periods of time the body is in the hyperglycemic state. As a result, patients often suffer from poor glycemic control, although whether or not it is related to obesity is strongly debated amongst the medical community. However, there are multiple studies that prove that poor glycemic control correlates with obesity, as shown in Figure 3 below (Gupta, 2020).

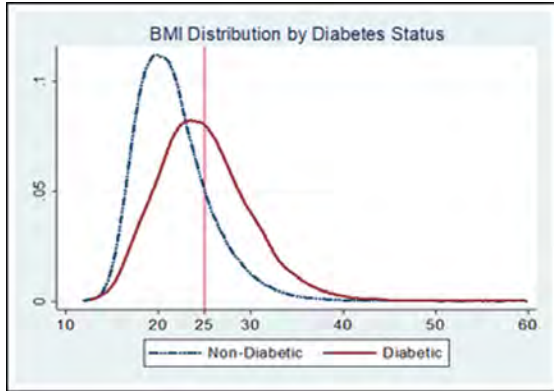


Figure 3. BMI Distribution by Diabetes Status. Figure constructed by author based on NFHS data for year 2015–16. (Gupta, 2020).

Diabetes treatment

Diabetes medicine has varied throughout the years; from injections to pills to even nasal sprays. The major component in all of these medications however is insulin. In both Type 1 (inherited) and Type 2 (developed) diabetes the pancreas (specifically the insulin-producing cells in the body) is negatively affected, limiting the body’s ability to produce its own insulin. As a result, scientists came up with an artificial pancreas and injectable insulin. Presently, in order for someone to treat themselves, they must prick their finger throughout the day to check their glucose levels and inject themselves with insulin to make up the difference needed (Figure 4)(Morris, 2021). Insulin can also be administered with an “artificial pancreas”, or an insulin pump.

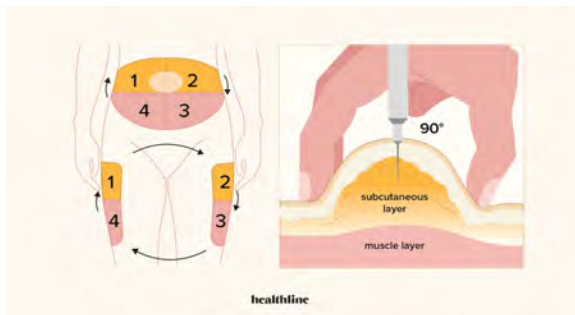


Figure 4. Diagram of Insulin Injection Technique. On the left is a diagram showing sites on the body where insulin injection should be targeted. The areas in yellow are areas high in fat, which better absorbs insulin than muscle (areas in red). On the right-hand side is a cross-section of the layers of skin, subcutaneous layer (fat) and muscle (Morris, 2021).

People with diabetes face many hardships, even with the vast amount of treatments out there. For example, a diabetic person must prick their finger up to eight times a day just to check their blood sugar. CGMs (continuous glucose monitors) were made to get rid of the pricks, however, users have to wear a patch on themselves every day, and still prick their finger to calibrate the machine. The patches are large and painful, and often require a large needle to insert the monitor (Figure 5)(Moses, 2019). We have found in this analysis that semaglutide will have a positive effect on patients, and could specifically help control blood glucose and reduce the risk of obesity. As previously explained, studies find that injection of semaglutide versus none at all promote more efficient weight loss results.

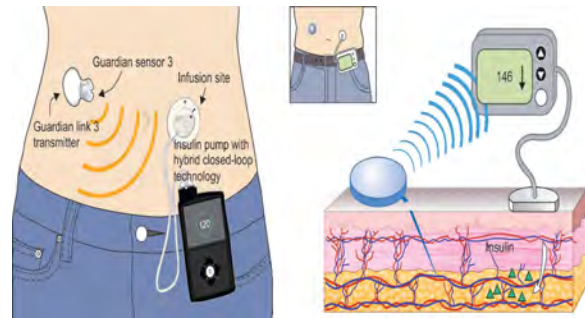


Figure 5. Diagram of a continuous glucose monitor and insertion. On the left is an example of a continuous glucose monitor along with an infusion site and a pump that connects to the glucose monitor. On the right is a diagram of how the glucose monitor communicates with the insulin pump (Moses, 2019).

Over the years, there have been very few drugs approved for the treatment of Diabetes. The most commonly recommended way to keep glucose levels under control is to exercise. Exercise can have a prominent effect on a person's body because when the body uses its muscles, it draws sugar from the blood to obtain energy. Depending on the duration or intensity of exercise, blood sugars can more or less be positively affected through physical activity (Roland, 2020). Exercise can improve peripheral insulin sensitivity as well as enhance insulin binding. Exercise also decreases abdominal fat, reduces free fatty acids, and increases insulin-sensitive skeletal muscle, which may result in improved glycemic

control (Whyte, 2013). Although it's been shown that in people with Type 2 diabetes exercise can improve insulin sensitivity (Whyte, 2013) this isn't always achievable or easily incorporated into regular habits by patients, or other chronic issues/injuries could prevent regular exercise.

Metformin, first approved in 1995 (Whyte, 2013), is another option for those who struggle with glycemic control. Although this drug is meant for those with Type 2 Diabetes, people with Type 1 Diabetes can also take and benefit from this medicine. Metformin works by helping to restore your body's proper response to the insulin you naturally produce. It also decreases the amount of sugar that your liver makes and that your stomach/intestines absorb (WebMD, Metformin).

Semaglutide, alongside weight loss, also treats unregulated glycemic control. Semaglutide can decrease appetite as well as help the pancreas to produce the right amount of insulin needed, thus leading to regulated glycemic control (Meier, 2021). The underlying mechanism for this drug involves targeting glucagon-like peptide-1 receptors. These are a well established class of glucose-lowering drugs that have been proven to help correct multiple issues that type 2 patients experience and can also help regulate glycemic control (Meier, 2021). As obesity decreases and glycemic control increases, this may lead, for example, a pre-diabetic person to become free from the potential risk of diabetes. As for a person with diabetes, this can greatly assist a person in their weight loss journey as well as keep them at healthy blood sugar levels. This could, in turn, mean fewer painful needle sticks throughout the day to monitor blood sugar levels.

2. Discussion

Despite significant progress in medical treatments for diabetes, the prevalence of obesity and type 2 diabetes remains high. This paper has explored the history of weight loss treatments, diabetes treatments, and the mechanisms of insulin and Semaglutide. Semaglutide, a new breakthrough in weight loss medication, has the potential to treat both obesity and type 2 diabetes. By examining current research on continuous glucose monitoring, glucose management,

insulin, and current treatment and preventative measures for diabetes, this paper has shown the relevance of Semaglutide in addressing these widespread diseases. These data indicate that semaglutide may promote more efficient weight loss in patients, and could specifically help control blood glucose and reduce the risk of obesity.

As medicine exponentially improves within the next few years, semaglutide may be a major benefactor to the field of diabetes. As a potentially effective weight loss drug, semaglutide can help save many who struggle with obesity and are on the brink of diabetes. Not only will this solve ongoing health issues related to the effects caused by obesity and diabetes, but it will also curb the number of cases of obesity per year by taking measures to help prevent those who are at risk of disease from acquiring it. This is only the first of many to come, and semaglutide will act as a reference to study from for years to come. Ongoing research studies and analysis aim to solidify the potential of Semaglutide as a treatment and prevention measure for obesity and type 2 diabetes. As scientists and researchers make further breakthroughs, the future of weight loss and diabetic control drugs will surely be bright.

Acknowledgments

I'd like to thank Heather Killeen Ph.D. for her feedback in preparing this manuscript for publication.

References

- Bansal, A. B., & Khalili, Y. A. (2022, July 21). *Orlistat*. NCBI Bookshelf.
- Boden, G. (2006). Fatty acid—induced inflammation and insulin resistance in skeletal muscle and liver. *Current Diabetes Reports*, 6(3), 177–181.
- CDC. (2022, March 2). *Type 2 diabetes*. Centers for Disease Control and Prevention.
- Cleveland Clinic. (n.d.). *Diabetes-Related ketoacidosis (DKA): Symptoms, treatment & causes*. Cleveland Clinic.
- Diabetes Teaching Center at the University of California, San Francisco. (n.d.). *Controlling blood sugar :: diabetes education online*. Diabetes

- Education Online; Diabetes Teaching Center at the University of California, San Francisco.
- Federal Trade Commission. (2014, September 11). *FTC settlement bans marketer behind 'fat burner' diet pills from manufacturing, marketing weight-loss products*. Federal Trade Commission.
- Felman, A. (2022, July 19). *What is hyperglycemia? Symptoms, treatments, causes, and all else you need to know*. Medical News Today.
- Ferraro, Z. M., Patterson, S., & Chaput, J.-P. (2015). Unhealthy weight control practices: Culprits and clinical recommendations. *Clinical Medicine Insights: Endocrinology and Diabetes*, 8, CMED.S23060.
- Gupta, S., & Bansal, S. (2020). Does a rise in BMI cause an increased risk of diabetes?: Evidence from India. *PLOS ONE*, 15(4), e0229716.
- Hughes, S., & Neumiller, J. J. (2020). *Oral Semaglutide*. Clinical diabetes: a publication of the American Diabetes Association, 38(1), 109–111.
- IBIS World. (2022). *IBISWorld - Industry market research, reports, and statistics*.
- Jato, J. A., Bawa, I., & Onyezili, F. N. (2018). Diabetes Induction with Streptozotocin and Insulin Action on Blood Glucose Levels in Albino Rats. *International Journal of Modern Science and Technology*, 3(10), 208–212.
- Johnson, D. B., & Quick, J. (2022, April 3). *Topiramate and phentermine*. NCBI Bookshelf.
- Klimek, P., Kautzky-Willer, A., Chmiel, A., Schiller-Frühwirth, I., & Thurner, S. (2015). Quantification of diabetes comorbidity risks across life using nation-wide big claims data. *PLOS Computational Biology*, 11(4), e1004125.
- Lily. (2019, January 1). *CGM Experiment: What I learned as a non-diabetic from wearing a continuous glucose monitor*. Lily Nichols RDN.
- Mayo Clinic. (n.d.). *Semaglutide (Subcutaneous Route) side effects*.
- Medical News Today. (2022, August 22). *How insulin and glucagon regulate blood sugar*. Medical News Today.
- Medline Plus. (n.d.). *Semaglutide injection: MedlinePlus drug information*. Medline Plus; National Library of Medicine.
- Meier, J. J. (2021). Efficacy of semaglutide in a subcutaneous and an oral formulation. *Frontiers in Endocrinology*, 12.
- Millar, H. (2021, March 30). *How does low blood sugar affect the body?* Medical News Today.
- Monaco, K. (2022, January 11). *Semaglutide vs liraglutide -- which induces more weight loss?* *MedpageToday*.
- Morris, S. Y. (2021, August 31). *Insulin injection sites: Where and how to inject*. Healthline Media.
- Moses, E. F. (2019). A hybrid closed-loop insulin delivery system for the treatment of type I diabetes. *Pondicherry Journal of Nursing*, 12(1), 25–27.
- NIDDK, National Institute of Diabetes and Digestive and Kidney Diseases. (2022, August 29). *Prescription medications to treat overweight & obesity*. NIDDK | National Institute of Diabetes and Digestive and Kidney Diseases.
- Padwal, R. S., et al. (2011). Using the Edmonton obesity staging system to predict mortality in a population-representative cohort of people with overweight and obesity. *Canadian Medical Association Journal*, 183(14), E1059–E1066.
- Prospective Studies Collaboration. (2009). Body-mass index and cause-specific mortality in 900 000 adults: Collaborative analyses of 57 prospective studies. *The Lancet*, 373(9669), 1083–1096.
- Roland, J. (2020, January 2). *How does eating affect your blood sugar?* Healthline Media.
- Sherman, M. M., Ungureanu, S., & Rey, J. A. (2016). Naltrexone/Bupropion ER (contrave): Newly approved treatment option for chronic weight management in obese adults. *Pharmacy and Therapeutics*, 41(3).
- Smith, M. (2022, August 16). *Facts about mealtime sugar spikes*. WebMD.
- Smith, M. W. (2014, April 3). *The Truth about Sensa*. WebMD.

Tsoukas, M. (2022, January 25). *Weekly subcutaneous semaglutide as adjunct to closed-loop therapy in type 1 diabetes care - Full text view*. ClinicalTrials.Gov; U.S. National Library of Medicine.

U.S. Food and Drug Administration. (2021, June 4). *FDA approves new drug treatment for chronic weight management, first since 2014*. U.S. Food and Drug Administration.

Victoza. (n.d.). *Possible side effects of Victoza*. Victoza® (Liraglutide) Injection 1.2 Mg or 1.8 Mg.

WCG CenterWatch. (n.d.-a). *Ozempic (semaglutide) subcutaneous injection*. CenterWatch.

WCG CenterWatch. (n.d.-b). *Victoza (liraglutide)*. CenterWatch.

WebMD. (n.d.). *Drugs & medications: Metformin HCl*. WebMD.

Whyte, J., & Mohr, C. (2013, January). *Exercise can improve glycemic control: How to get patients started*. Consultant360.

Wright, S. A. (2022, May 25). *What to know about weight loss patches*. Healthline Media.

Zhao X, et al. (2021) *GLP-1 Receptor Agonists: Beyond Their Pancreatic Effects*. Front. Endocrinol. 12:721135. doi: 10.3389/fendo.2021.721135.

Code-switching: An Analysis of Causes, Benefits, and Perceptions

Carey Chang^{1*}

¹The Harker School, San Jose, CA USA

Received February 7, 2023; Revised April 17, 2023; Accepted, May 9, 2023

Abstract

Code-switching is the linguistic phenomenon of alternating between two languages or dialects during a conversation. This paper explores the various factors that influence the practice of code-switching, focusing on those that are sociolinguistic, conversational, or a combination of both. By examining these factors, this paper aims to shed light on the complexities of code-switching and how it reflects larger societal attitudes toward language and identity. Through these factors, the paper sets to prove that code-switching is natural and predictable, making it something not to be discouraged or looked down upon as many teachers or parents believe. As such, this paper also addresses the issue of prescription versus description in code-switching research, arguing that linguists should adopt a descriptive approach to language, rather than a prescriptive one.

Keywords: Code-switching, Multilingualism, Prescription vs. Description, Language identity, Discourse analysis,

1. Introduction

Utilized by countless individuals across the world, code-switching is the changing of one language or dialect to another during conversation. This fascinating phenomenon is not random, but rather follows a set of specific linguistic constraints (Poplack, 1980). One example is the equivalence constraint, which posits that a code-switch takes place such that it “does not violate a syntactic rule of either language” (p. 586). Another is the free morpheme constraint, which states that “codes may be switched after any constituent in discourse provided that constituent is not a bound morpheme” (p. 585-6). For instance, “malinterpreté” in Spanish means “misunderstood”. A code-switch such as “misinterpreté” is not possible as “mal-/mis-” is a bound morpheme. Studying code-switching allows linguists to better comprehend the complexities of the language and psychological processes of multilingual

humans (Beatty-Martínez, et al., 2020), which is crucial considering the majority of the world speaks more than one language.

Even though code-switching is often an instinctive process that occurs subconsciously in conversation (MasterClass, 2022), historically there has been a stigma against code-switching. From personal experience growing up as a Taiwanese-American, parents often discourage their children from code-switching to English while having conversations in Mandarin with them. Like many others, they believe that code-switching negatively affects language ability and that the capability of those who code-switch is worse than those who use the proper form (Kight, 2018). However, it has been found that code-switching actually demonstrates proficiency in the non-base language (Rouchdy, 2003). In Korea, Korean English, also known as Konglish, is considered a failed attempt at proper English. Opponents of

* Corresponding Author
24careyc@students.harker.org

Advisor: Dr. Ekarina Winarto
ew385@cornell.edu

code-switching in Korea judge Konglish based on standard English's grammatical rules and believe it to be inferior (McPhail, 2018). However, Konglish is neither Korean nor English and should not be criticized based on the rules of either language.

To dispel the misconceptions surrounding code-switching, it is important to first understand the different factors that influence this phenomenon. This paper aims to show that code-switching is a natural and beneficial element of speech, by discussing the various uses and advantages of code-switching through an investigation of why people code-switch in the first place. The first section of the essay will analyze sociolinguistic factors that are present independent of the conversation, which are associated with entire populations or communities, rather than affecting a particular conversation. Next, this paper will cover conversational factors which, unlike the sociolinguistic factors, are distinct occurrences within the discourse itself that directly cause code-switching. Lastly, this paper discusses factors that are a combination of both sociolinguistic and conversational factors, which impact specific parts of conversation but are influenced by larger societal circumstances. Due to its benefits, this paper proposes that linguists should not prescribe how people use code-switching, but rather describe code-switching as it is used in spontaneous speech, as a prescription is useless so long as code-switching is widely used in conversation. Prescription can be damaging to a bilingual speaker's ability to communicate and sense of identity, so it is important to erase the stigma surrounding code-switching.

2. Sociolinguistic Factors

One main cause that affects the use of code-switching is those caused by someone's socioeconomic status, known as sociolinguistic factors. One such factor is the desire to preserve one's culture. An example of this is the Turkish-speaking population in Greece, who have been observed to intentionally abstain from code-switching to Greek, due to a "high level of awareness of the need to protect their language and culture from Greek influence" (Gardner-Chloros, 2012, p. 104). Additionally, the community in which the

code-switching occurs also affects the different types of code-switching that appear. For instance, knowledge of English is prevalent in both Spain and Mexico, but code-switching has a wider range of uses in Mexico, where it "is used in ironic contexts to reflect a certain rejection of the US culture" (p. 102). Code-switching can also be used to help speakers better assimilate within a group. Linguist Richard Bailey found that Dominican-American youth in Providence, Rhode Island did not ethnically identify as White or Black (Nilep, 2006, p. 13). Their intricate identity was shown by their use of nonstandard English varieties such as Dominican Spanish, Caribbean Spanish, and African American Vernacular English (p. 14). In a conversation, code-switching acts as a "political strategy" (p. 12), allowing these individuals to "to say and do, indeed to be two or more things where normally a choice is required," (Heller, 1988, p. 93) as language is often perceived as closely related to one's identity. Their background permits them to choose different dialects and languages to match the appropriate situation, making their speech more versatile and adaptable.

Other than changing between dialects, code-switching involving the vocabulary used in a conversation, such as using slang, can also form ingroups and outgroups. Just like technical jargon, which differentiates people of different professions such as linguists and economists, slang also separates individuals into those who understand and those who do not (Mazrui, 1995). There are numerous examples of slang in the English language, often used by the younger generation and spread through the internet. The phrase "spill the tea," popularized in the early 2010s, literally means to overturn a container holding a warm liquid with tea leaves; however, it is now more commonly used to mean "share the gossip." Although the slang in this example does not code-switch to a different language, the switch from standard English to an informal expression is a change in the "discourse type" (Chandler & Munday, 2011). Using slang creates a feeling of unity in a group, as shared knowledge of specific terms and references makes participants feel more involved (Mazrui, 1995). While slang socially dissociates the speaker from the outgroup, it also reinforces unity among the ingroup by allowing them to form a

separate, shared identity (p. 172).

3. Conversational Factors

Additionally, some factors that affect code-switching are purely conversational, which arise within the dialogue, independent of the socioeconomic status of the speaker. One such factor is during quotation, when a speaker refers to a previous situation or other speakers during a conversation. In McClure & McClure (1988), a Saxon woman describes her encounter with a Romanian policeman: “ən do vor əzi ə gaŋ məlitsyán ən zot ke mix^y : *Aici e pașaporta, acumă poți să pleci în Germania,*” which translates to “And there was a young policeman and he said to me: *Here is your passport, now you can leave for Germany.*” In this situation, the code-switching benefits the conversation because the policeman’s words are most accurately represented in Romanian (the italicized portion), as that was the language spoken in that conversation. If the woman instead spoke completely in Saxon, the translated meaning or tone of what the police officer had said may have been slightly altered from what had been intended. In certain situations, such as this one which involves police, it is especially important for speech to be repeated verbatim so as to not cause misunderstandings.

Another reason why bilinguals use code-switching is to emphasize certain emotions or phrases in a conversation. For instance, Lantto (2014) observed that Basque-Spanish speakers in Great Bilbao, a municipality in Spain, code-switch to Spanish while cursing when Basque is the base language of the conversation. The paper provided many possible reasons why this occurred. One such explanation was to accentuate the swearing by creating a sharper contrast between the curse words and the rest of the conversation. This emphasizes the feelings of emotion that the speaker is experiencing by drawing more attention to the swearing, therefore enhancing the overall conversation. If the cursing had been in Basque instead, its effect would have been more muted.

Along with swearing, speakers can also use interjections to express sudden emotion and draw attention to that emotion. Numerous examples of

interjections exist in English, such as ‘Wow!’, ‘Yay!’, and ‘Ouch!’ An example is in Gumperz (1982), with a Bahasa Indonesia - English Bilingual saying, “*Oh dear, kamu baik-baik saja kan? Gak ada yang cedera kan?*” (p. 77). This translates to “Oh dear, you are okay, right? Nobody got hurt, right?” The English interjection draws the attention of the listener to the speaker, as the code-switching is a distinct difference from the rest of the sentence, spoken in Indonesian. Additionally, it emphasizes the worry and strong emotion that the speaker is feeling (Tamara, 2018).

To ensure successful communication, those involved often have to cooperate. In code-switching, this could be done by copying code-switching patterns used by a conversation partner, also known as priming. By observing Spanish-English bilinguals, Fricke & Kootstra (2016) suggests that a conversation partner is more likely to code-switch if the other partner recently code-switched in the conversation (up to 10 sentences prior). Similarly, Korean-English bilingual children have also been recorded to change their use of Korean or English based on the language choice of others. If one child signals a preference to speak English, then other children are more likely to utilize English to communicate with that child (Shin & Milroy, 2000). By choosing the optimal combination of Korean and English that suits both parties, the schoolchildren are able to communicate most effectively.

4. Sociolinguistic and Conversational Factors

Although some factors are purely sociolinguistic or conversational, oftentimes factors are a combination of both. These are the factors that influence specific linguistic elements of a conversation, but are still impacted by the speaker’s overall values. As language and culture are closely related, a country’s culture is often reflected in its language’s vocabulary. For instance, emotion words tend to be more intense in languages whose cultures are considered “warmer.” In a study of English/Greek bilinguals, Panayiotou (2004) reports that participants would often code-switch to that language if specific terms to describe the participants’ feelings were more suitable in one language over the other. In particular, the volunteers chose to use English to express their

emotions with words such as “indifferent,” “concerned,” “frustrated,” and “sympathy” (p. 132). Oftentimes, the responders code-switched to English, because there was no Greek word to represent the meaning they wished to convey and the English words were viewed as less emotional, thereby conveying their intended meaning better. The participants were able to express emotions fluently in both languages, but they chose to code-switch to English in order to make their ideas clearer, not because of a preference for either language. The findings from this study suggest that there are certain experiences that can only be properly expressed in one language and that translation into another language causes a loss of cultural significance and context (p. 133). Code-switching allows these bilinguals to better express their feelings, as their experience may be common in one culture but nonexistent in another, making one language significantly more suitable for representing their situation.

Similarly, some words or phrases carry more meaning in certain languages than in others, where they have a cultural significance that surpasses the literal meaning of the words. For example, a phrase that serves of a personal reminder of childhood memories and traditions when it is spoken is “新年快乐 (xīn nián kuài lè),” which translates to “Happy Chinese New Year” in English. Growing up, “新年快乐” was associated with spending time with friends and families, seeing bright red and gold decorations, and receiving money in red envelopes. On the other hand, the phrase “Happy Chinese New Year” does not evoke the same feeling, as it does not have the same cultural connections in English. Thus, the choice to use certain terms in a specific language can enhance the meaning due to the social roots that are linked to that language.

Additionally, code-switching can be used for comedic effects. One such example is translinguistic puns, which take advantage of similarities between words or phrases across two languages (Caubet, 2003). Mohamed Fellag, an Algerian comedian, commonly uses these puns. For example, he references Wall Street as “*Waal* Street,” meaning “*Nothing* Street” (p. 247). This pun plays on the phonological similarity between the Arabic word

“*waalu*,” meaning “nothing,” and the English word “wall.” The use of two languages in this example, English and Arabic, allows for jokes that would be otherwise unable to be expressed. This type of humor is possible in part due to the diversity of Fellag’s audience, which is known to be multicultural and multilingual, stemming from outside influences as a result of colonization throughout the history of that region. Comedians such as Fellag use code-switching to leverage this background for comedic effect as well as to comment on the “social or political situations in their countries” (p. 253).

5. Description versus Prescription

Due to its various functions, code-switching is important in interpersonal conversation between bilinguals. Each factor, whether it leads to conscious or subconscious code-switching, improves the conversation by building closer connections with others, elucidating the speaker’s intent, or drawing attention to the speaker. Therefore, code-switching should be viewed as yet another tool for multilinguals to most effectively communicate, just like gestures, facial expressions, tone, intonation, and volume. Because of its plethora of uses, it will continue to be practiced even if linguists do prescribe whether code-switching is good or bad. As long as code-switching retains a purpose for speakers, such a prescription would be functionally useless, as there is no value in a prescription that does not represent the behavior of the general population. Furthermore, prescribing code-switching rules at all is difficult due to its situational adaptability. As each multilingual uses code-switching and language in unique ways, there are countless possibilities for code-switching to be employed, with no one way superior to another. For example, when individuals use slang, as described above, they may be deviating from what is considered “standard” language. However, slang has its own benefits, and it is unreasonable to judge that slang is inferior to proper language. Instead, it is more important to describe code-switching as it manifests in conversation, as this description can provide insights into the languages involved. As previously stated, Panayiotou (2004) found that English/Greek bilinguals switched to English to

describe feelings that were less emotional. Similar results with other languages could provide information about the perceived “warmness” of those languages.

6. Conclusion

Although there has been a bias against code-switching, active research is currently being performed to clarify the misconceptions that have stigmatized it. By understanding the factors of code-switching, such as those that are sociolinguistic and/or conversational, linguists can disprove assumptions such as the notion that code-switching is “cheating” (Yuhua, 2021). This work has even allowed the study of code-switching to impact disciplines beyond linguistics. Moreover, studying code-switching has allowed for improvements in these additional subject areas. In the field of education, translanguaging, a newer concept of using multiple languages during learning, is gaining traction (EAL, 2016). By eliminating the rigid one-language learning model in certain situations, students can take advantage of their “full linguistic repertoires” (Hamman, et al., 2018, para. 11) to better grasp concepts and learn more efficiently. Through advancements such as translanguaging, code-switching is becoming more common and accepted, improving the communication and lives of multilinguals across the world.

Acknowledgments

I would like to express my gratitude toward Dr. Winarto for her advice and support throughout the entire research process. Her help was crucial for this paper.

References

Beatty-Martínez, A. L., Navarro-Torres, C. A., & Dussias, P. E. (2020). Codeswitching: A bilingual toolkit for opportunistic speech planning. *Frontiers in Psychology, 11*.
<https://doi.org/10.3389/fpsyg.2020.01699>

Chandler, D., & Munday, R. (2011). A dictionary of media and communication.

<https://doi.org/10.1093/acref/9780199568758.001.0001>

EAL Writers. (2016, July 26). *What is translanguaging?* EAL Journal. Retrieved July 26, 2022, from
<https://ealjournal.org/2016/07/26/what-is-translanguaging/>

Fricke, M., & Kootstra, G. J. (2016). Primed codeswitching in spontaneous bilingual dialogue. *Journal of Memory and Language, 91*, 181–201.
<https://doi.org/10.1016/j.jml.2016.04.003>

Gardner-Chloros, P. (2009). Sociolinguistic factors in code-switching. *The Cambridge Handbook of Linguistic Code-Switching*, 97–113.
<https://doi.org/10.1017/cbo9780511576331.007>

Gumperz, J. J. (1982). *Discourse strategies*. Cambridge Univ. Pr.

Hamman, L., Beck, E., & Donaldson, A. (2021, May 18). *A pedagogy of translanguaging*. Language Magazine. Retrieved July 17, 2022, from
<https://www.languagemagazine.com/2018/09/10/a-pedagogy-of-translanguaging/>

Heller, M. (1988). *Codeswitching anthropological and sociolinguistic perspectives*. Mouton de Gruyter.

Kight, B. (2018, September 16). *The Code-Switch Stigma*. Thebottomlinenews.com. Retrieved July 17, 2022, from
<https://thebottomlinenews.com/the-code-switch-stigma/#:~:text=Code%2Dswitching%20can%20be%20defined,in%20certain%20families%20or%20cultures.>

Lantto, H. (2012). Code-switching, swearing and slang: The Colloquial Register of Basque in Greater Bilbao. *International Journal of Bilingualism, 18*(6), 633–648. <https://doi.org/10.1177/1367006912457274>

MasterClass, M. C. S. (2022, January 5). *Code-switching definition: 5 reasons people code-switch - 2022*. MasterClass. Retrieved July 15, 2022, from
<https://www.masterclass.com/articles/code-switching-5-functions-of-codeswitching>

Mazrui, A. M. (1995). Slang and code-switching: the case of Sheng in Kenya.

McPhail, S. A. (2017). South Korea's linguistic tangle: English vs. Korean vs. Konglish. *English Today*, 34(1), 45–51.
<https://doi.org/10.1017/s0266078417000244>

Nilep, C. (2006). "Code Switching" in Sociocultural Linguistics. *Colorado Research in Linguistics*.

Panayiotou, A. (2004). Switching codes, switching code: Bilinguals' emotional responses in English and Greek. *Journal of Multilingual and Multicultural Development*, 25(2-3), 124–139.
<https://doi.org/10.1080/01434630408666525>

Poplack, S. (1980). Sometimes I'll start a sentence in Spanish y Termino en español: Toward a typology of code-SWITCHING1. *Linguistics*, 18(7-8).
<https://doi.org/10.1515/ling.1980.18.7-8.581>

Rouchdy, A. (2002). *Language contact and language conflict in Arabic: Variations on a sociolinguistic theme*. Routledge.

Shin, S. J., & Milroy, L. (2000). Conversational codeswitching among Korean-english bilingual children. *International Journal of Bilingualism*, 4(3), 351–383.
<https://doi.org/10.1177/13670069000040030401>

Tamara, A. F. (2018). Types and Speech Functions of Using Code Switching by the Main Character in Critical Eleven Novel.

Yuhua, D. (2021, November 30). *How brains seamlessly switch between languages*. Scientific American. Retrieved June 29, 2022, from <https://www.scientificamerican.com/article/how-brain-s-seamlessly-switch-between-languages/>

Generalized Anxiety Disorder and Sleep Quality during the COVID-19 Outbreak in Adolescents & Parents: A Web-Based Cross-Sectional Survey

Kaylin Spinelli¹*, Stephen J. Sullivan¹

¹Sacred Heart Academy, Hempstead, NY USA

Received February 17, 2022; Revised March 18, 2023; Accepted, April 3, 2023

Abstract

Many previous studies have shown that one's anxiety is often heightened during disease pandemics. The present study found evidence that essential workers and their families are at an increased risk for anxiety, stress and sleep disruption. 160 female adolescents and 75 of their parents completed surveys measuring anxiety and sleep quality during the COVID-19 pandemic. The two samples, although drawn from the same community, were not linked. All adolescents attend an all-girls academy in suburban New York. T-tests and regression analyses suggest that essential workers and their families are at elevated risk for mental illness.

Keywords: Anxiety, Mental health, COVID-19, Frontline workers, Sleep quality

1. Introduction

The stress levels of countless people throughout the world have grown exponentially during the COVID-19 pandemic. Numerous studies show that one's anxiety is often heightened during disease outbreaks (Tausczik, et al., 2012; Grover, et al., 2020). Studies also show that lack of mental health help during disease outbreaks can play a role in the spread of the disease and in the emotional stress during and after the outbreak. (Cullen, et al., 2020). Every individual copes with stress and anxiety differently, they are brought on in a variety of ways for people at different times. Some individuals cope with anxiety and stress by sleeping more, while others find it difficult to fall asleep at all during stressful or anxious times. The severity of the COVID-19 pandemic has certainly increased anxiety more significantly in more people, than most other diseases. Another reason is because people have been confined to their homes for extended periods of time

(Pfefferbaum & North, 2020). This pandemic was especially stressful because it was particularly hard to control things in larger cities. (The Lancet, 2020; Ciotti, et al., 2020). "Implementing public health measures is difficult in places with overcrowded living conditions and inadequate hygiene and sanitation." (The Lancet, 2020) After almost two years after "patient zero" positive tests and quarantines remain too common.

Teens are developmentally prone to high anxiety levels (Henker, et al., 2002). At their age, it is most common to worry about peer acceptance and academic performance (Henker, et al., 2002). Now teens feel they have additional anxiety brought on by a once-in-a-century pandemic. Gupta, et al., (2020) found that during the lockdown phase of the pandemic, many individuals reduced their hours of nighttime sleep-in favor of taking more mid-day naps. This pattern slowly undermined mental health (Gupta, et al., 2020). Lee, et al., (2020) found that the two most common mental health disorders to come

* Corresponding Author
kkspinelli164@gmail.com

Advisor: Stephen J. Sullivan (Ph.D.)
sully49@gmail.com

from a pandemic are post- traumatic stress disorder and anxiety.

A study of healthcare workers in Sudan found heightened anxiety during the COVID-19 pandemic (Elamin, 2020). This raises the question of the extent to which the heightened anxiety levels of healthcare workers might affect their families. El-Hage, *et al.*, (2020) states that many healthcare workers have suffered from mental health issues during the pandemic. Resources to promote mental well-being in health care professionals needs to be a standard part of disease outbreak protocol (El-Hage, *et al.*, 2020). The present study examines not only the anxiety levels and sleep quality of individuals who self-identify as essential workers, but also those individuals who self-identify as the child of an essential worker. Essential workers were classified people who worked on the front lines during the pandemic and could not work from home. A few examples are Doctors, nurses, paramedics, etc. Did these two groups report significantly higher COVID-19 related anxiety or more significantly disrupted sleep than peers from a non-essential worker family?

Our study is a cross-sectional, web-based survey testing generalized anxiety disorder and sleep quality during the COVID-19 pandemic in adolescents and their parents. A similar study of adult workers - frontline and non-frontline - was conducted in China in March 2020 at the height of the pandemic. We intentionally followed similar research protocols to better compare our February 2021 results in New York with the February 2020 data collected by Chinese scholars. Huang & Zhao (2020) found that younger female people and essential workers were at highest risk of mental illness due to the COVID-19 outbreak. The present research tests whether the mid-pandemic anxiety levels and sleep quality of essential workers and their children differed from that of their peers using the *Generalized Anxiety Disorder Survey (GAD-7)* and the *Pittsburgh Sleep Quality Index questionnaire (PSQI)*. The *GAD-7* is a valid and efficient tool for screening for Generalized Anxiety Disorder and assessing its severity in clinical practice as well as in research.

The aim of this study is to (1) gain a better understanding of the effects of a global pandemic on

adolescents and their parents and, (2) to uncover the impact essential workers have on the anxiety levels and sleep quality of their families during a pandemic - present or future. It is hypothesized that essential workers and their family members will be at higher risk for anxiety and suffer from poorer sleep quality. It is also hypothesized that adolescents will report greater anxiety levels (higher *GAD-7* scores) and poorer sleep quality (as per *PSQI*) compared to adults.

2. Methods and Materials

2.1 Participants

We recruited 160 female adolescents from an all-girls Catholic academy in a suburb of New York City and 75 of their parents. However, students' and parents' respective surveys were not linked. We observed both parents and children but did not look for direct comparison between parents and their own child(ren). We sought to survey adults and teens from the same community, not necessarily test for specific parental influence on individual girls. The participants were told that the project involved "sleep quality and anxiety levels during the COVID-19 pandemic." The sample was reasonably representative of the racial/ethnic makeup of the school. The school, although "increasingly diverse," is more Euro-American in makeup than the surrounding county (see Table 1).

2.2 Procedure

After informed consent was received from the adolescent subjects and one custodial parent/guardian (study 1) or from adult subjects (study 2). Participants completed the survey electronically via Google Forms. Students were identified by their school-issued ID numbers', parents were identified by the last four digits of their phone number and their birth month. Although all responses were kept anonymous, researchers sought to discreetly identify subjects in case respondents requested that their information be removed from the study at a later date. Each subject completed demographic information, questions about parents/guardians' schooling and

profession (or their own schooling and profession if adults were the subjects). Subjects next took the Generalized Anxiety Disorder-7 (*GAD-7*) survey, and the Pittsburgh Sleep Quality Index (*PSQ*) questionnaire. The *GAD-7* has been proven to be one of the most effective methods of measuring anxiety (Spitzer, *et al.* 2006). Parents were classified as

essential or non-essential workers. Thus, two measures of “essential worker” were used - the respondent’s and New York States’. From there we were able to test how the jobs of parents might affect the anxiety levels and the sleep quality of their children.

Table 1: High School Ethnicity Statistics

Ethnicity*	High School Population**	Nassau County ⁺	United States ⁺	Sample % for adolescents	Sample % for parents
White	74%	58.5%	60.1%	70.2%	64.3%
Black	10%	13.1%	13.4%	12.9%	14.3%
Asian	11%	10.9%	5.9%	5.6%	3.6%
Hispanic	4%	17.5%	18.5%	11.3%	17.9%
Other	-	2.6%	4.3%	1.6%	0%

* Totals may not equal 100 due to those who reported themselves as multiracial.

**2018-19 data provided by Sacred Heart Academy Annual 2019 Report to NYSED

+ 2019 U.S. Census Estimate

Table 2: Sample Gender Statistics

Gender*	Sample % for adolescents	Sample % for parents
Male	0%	38%
Female	100%	62%

* The high school population studied was drawn from an all-girls academy.

Table 2.5: Sample Age Statistics

Age*	High school population	Sample % for adolescents
13	1.0%	0%
14	25.7%	16.1%
15	22.8%	25.8%
16	31.7%	27.4%
17	13.9%	25.8%
18	4.9%	4.8%

*This sample consists of 120 high school students from an increasingly diverse Catholic all-girls high school

Table 4: Anxiety and Sleep Quality Scales

Title	Author(s) / Date	Purpose / Sample Questions
Generalized Anxiety Disorder 7-Item (<i>GAD-7</i>) Scale	Spitzer, <i>et al.</i> (2006)	Over the last 2 weeks, how often have you been bothered by the following problems? (Likert scale of 0 to 3) #1 Feeling nervous, anxious, or on edge. #2 Not being able to stop or control worrying. #3 Worrying too much about different things.
Pittsburgh Sleep Quality Index (<i>PSQI</i>)	Buysse, <i>et al.</i> (2010)	#1 During the past month, what time have you usually gone to bed at night? #2 During the past month, how long (in minutes) has it usually taken you to fall asleep each night? #3 During the past month, what time have you usually gotten up in the morning?

Table 5: Variables

Independent (x)	Dependent (y)	Covariates
Essential worker - vs. - Nonessential worker	GAD-7 score	Age, race/ethnicity, SES, previous mental health condition
Child of essential worker - vs. - Child of nonessential worker	PSQI sleep hours subscore	

3. Results

Teens reported a significantly higher prevalence of anxiety symptoms (*GAD-7*) than older people. However, adults did not report achieving significantly more sleep, measured in hours. Occupationally, workers self-identifying as essential workers and teens who self-identified themselves as the children of essential workers were significantly more likely to report poor sleep quality (all p 's < .05), and to report higher anxiety levels (p 's < .05). A series of multiple regressions demonstrated that age group (teen vs. adult) and occupational status were predictive of both generalized anxiety (*GAD-7*) and poor sleep quality (*PSQI*); all p 's < .05. However, age group dropped out of the multiple regression model (both p 's > .05); occupational status and the subsequent perceived risk was the only direct predictor of both anxiety and poor sleep (both p 's < .05). As hypothesized, essential workers and their family members were proven to be at higher risk for anxiety; essential workers and their adolescent family members were also found to have poorer sleep quality. The adolescents proved to have greater anxiety levels (higher *GAD-7* scores) and poorer sleep quality (as per *PSQI*) compared to adults. These findings can also be found below in Tables 6 and 7.

Table 6: T-tests - Means & Standard Deviations

	GAD-7*	hrs/sleep*
Teens	10.75 ^a (1.39)	6.90 ^a (1.47)
Adults	9.44 ^b (1.29)	6.50 ^b (1.38)

*Means with differing superscripts differ at the 95% confidence level.

Sample Regressions for *GAD- 7* scale

$$r^2_{\text{teen} \times \text{adult}} = .072 \mid 7.2\% \text{ of the variance in anxiety}$$

could be accounted for by age group alone.

$$r^2_{\text{FW} \times \text{N-FW}} = .169 \mid 16.9\% \text{ of the variance in anxiety could be accounted for by occupational status (essential worker/child of vs. non-essential worker/child of)}$$

Sample Regressions for *PSQI* sleep scale

$$r^2_{\text{teen} \times \text{adult}} = .039 \mid 3.9\% \text{ of the variance in sleep hours could be accounted for by age group alone.}$$

$$r^2_{\text{FW} \times \text{N-FW}} = .199 \mid 19.9\% \text{ of the variance in sleep hours could be accounted for by occupational status (essential worker/child of vs. non-essential worker/child of)}$$

Table 7: T-tests - Means & Standard Deviations

	GAD-7*	hrs/sleep*
Teens _{FWparent}	11.10 ^a (1.31)	7.33 ^a (1.08)
Teens _{N-FWparent}	9.35 ^b (1.43)	6.45 ^b (1.34)
Adult _{FW}	6.84 ^a (1.25)	6.95 ^a (1.12)
Adult _{N-FW}	6.02 ^b (1.34)	6.08 ^b (1.12)

*Means with differing superscripts differ at the 95% confidence level.

4. Discussion and Limitations

Our study identified a significant behavioral health burden on the American public nine months into the COVID-19 outbreak. 1) Younger people and 2) essential workers and their adolescent family members appear to be at elevated risk for mental illness, and may be in need of behavioral intervention. Our results were in line with those of Tausczik, *et al.* (2012), in that we both found that one's anxiety level often becomes heightened during disease outbreaks. Ongoing surveillance of the psychological consequences of disease outbreaks must become a standard part of preparedness

protocols in the United States and worldwide. Our results also supported the findings of Pfefferbaum and North, (2020); both studies found that the pandemic increased anxiety. Our results were in line with those of the study conducted in China back in March of 2020 (Huang & Zhao, 2020). One of the limitations of our study is that thus far we have only examined self-reported essential worker status. These results were also entirely quantitative in nature. To extend this research, it might be wise to include a qualitative study involving focus groups of a subsample of our survey respondents. Discussion would revolve around the reasons behind their anxiety - specific triggers, preferred coping strategies, etc. This would create a more comprehensive mixed-method study that could answer additional questions. Another limitation of this study is the sample size. Our sample of parents remains smaller than that of our adolescent participants. Mid-pandemic, recruiting parents to take even an online survey was difficult because we lacked any in-person connections to the parents. Standard mother-daughter, father-daughter, and parent-teacher events at school were canceled or moved to virtual platforms. Lastly, another limitation of this study is that there are many other factors that may have impacted worsened anxiety and sleep quality during the pandemic such as transition to online school or differences in lifestyle and routine. In sum, this study has generally supported my hypotheses, but also brings to light new questions worthy of study in further research.

Acknowledgements

I would like to thank Dr. Stephen Sullivan, my research advisor, for inspiring my interest in behavioral research and for closely reading several drafts of every abstract, paper and PowerPoint. I would also like to thank Dr. Beth Feinman, the head of the science department at my school, for serving as an excellent role model of a professional female scientist. Lastly, I want to thank my parents for always supporting me in everything I do but especially for their support throughout my research.

References

- Casella, M., *et al.* (2021). Features, Evaluation, and Treatment of Coronavirus (COVID-19). In *StatPearls*. StatPearls Publishing.
- Ciotti, M., *et al.* (2020). The COVID-19 pandemic. *Critical Reviews in Clinical Laboratory Sciences*, 57(6), 365–388.
- Cullen, W., *et al.* (2020). Mental health in the COVID-19 pandemic. *QJM : Monthly Journal of the Association of Physicians*, 113(5), 311–312.
- Elamin, M. (2020). The Psychological Impact of the COVID-19 Pandemic on health professionals in Sudan 2020. *Sudan Journal of Medical Sciences (SJMS)*, 15(5), 54-70.
- El-Hage, W., *et al.* (2020). Health professionals facing the coronavirus disease 2019 (COVID-19) pandemic: What are the mental health risks? *L'Encephale*, 46(3S), S73–S80.
- Grover, S., *et al.* (2020). Psychological impact of COVID-19 lockdown: An online survey from India. *Indian Journal of Psychiatry*, 62(4), 354–362.
- Gupta, R., *et al.* (2020). Changes in sleep pattern and sleep quality during COVID-19 lockdown. *Indian journal of psychiatry*, 62(4), 370–378.
- Henker, B., *et al.* (2002). Anxiety, affect, and activity in teenagers: monitoring daily life with electronic diaries. *Journal of the American Academy of Child and Adolescent Psychiatry*, 41(6), 660–670.
- Huang, Y. & Zhao, N. (2020). Generalized anxiety disorder, depressive symptoms and sleep quality during COVID-19 outbreak in China: a web-based cross-sectional survey. *Psychiatry Research*, 288.
- Javelot, H., & Weiner, L. (2020). Panic and pandemic: Review of the literature on the links between panic disorder and the SARS-CoV-2 epidemic. *L'Encephale*, 46(3S), S93–S98.
- Lee, S., *et al.* (2020). Incremental validity of coronaphobia: Coronavirus anxiety explains depression, generalized anxiety, and death anxiety. *Journal of Anxiety Disorders*, 74.

Lee, S., Jobe, M. C., & Mathis, A. A. (2020). Mental health characteristics associated with dysfunctional coronavirus anxiety. *Psychological Medicine*, 1–2. Advance online publication.

Lee, S., *et al.* (2020). Clinically significant fear and anxiety of COVID-19: A psychometric examination of the Coronavirus Anxiety Scale. *Psychiatry Research*, 290.

Pfefferbaum, B. & North, C. (2020). Mental Health and the Covid-19 Pandemic. *The New England journal of medicine*, 383(6), 510–512.

Spitzer, R., *et al.* (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of internal medicine*, 166(10), 1092–1097.

Tausczik, Y., *et al.* (2012). Public anxiety and information seeking following the H1N1 outbreak: blogs, newspaper articles, and Wikipedia visits. *Health communication*, 27(2), 179–185.

The Lancet (2020). India under COVID-19 lockdown. *Lancet (London, England)*, 395, 1315.

Influence of pH on Coordination Reactions of Humic Acid with Metal Ions

Keilah Om^{1*}

¹Dominican Academy, New York, NY USA

Received September 10, 2022; Revised March 19, 2023; Accepted, May 1, 2023

Abstract

The coordination reactions of Humic acid with metal ions were studied. Using the ion exchange equilibrium method, we can find the stability constants of complexes ($\log K$) and the coordination number (x) of these reactions. The results showed that in the reactions of HA-Fe³⁺, HA-Zn²⁺, and HA-Cd²⁺, the stability constants of complexes and coordination numbers increased as the pH increased and when ionic strength and temperature conditions were kept the same. For the reaction of HA-Mn²⁺, the stability constants of complexes increased as the pH increased, and the coordination number increased as the pH increased during the range of 3.0 to 5.0; however, it decreased when the pH was increased from 5.0 to 7.0.

Keywords: pH, Coordination Reaction, Humic acid, Fe³⁺, Mn²⁺, Cd²⁺, Zn²⁺

1. Introduction

1.1 Humic Acid

Humic acid is a principal component of Humic substances, which are the major organic constituents of soil (humus), peat, coal, many upland streams, dystrophic lakes, and ocean water (Stevenson, 1994). Humic consists of 40-60% carbon, 30-50% oxygen, 1-3% nitrogen, and 0.1-2% sulfur combined in aliphatic and aromatic ring structures associated with carbonyls, alcoholic and phenolic hydroxyl, carboxylates, amines, amides, and other functional groups (Baker, et al., 2007).

Humic acid is the most widely encountered natural complexing ligand, which tends to chelate various metal ions present in the environment. In Humic acid-Metal complexes, Humic acid can change the properties of metal ions. For instance, metal solubility can be increased when metals are complexed with free Humic acid, or decreased when

metals are scavenged by Humic films on mineral surfaces (Zhou, et al., 2015)

The presence of carboxylate and phenolate groups gives the Humic acids the ability to form complexes with ions. Many Humic acids have two or more of these groups arranged so as to enable the formation of chelate complexes (Tipping, 1994). The formation of (chelate) complexes is an important aspect of the biological role of Humic acids in regulating bioavailability of metal ions (Ghabbour, et al., 2001).

1.2 Measurement principle of stability constants of complexes ($\log K$) and the coordination number (x)

A known weight of a well-defined ion exchange resin is added to a solution containing a mixture of a Humic acid and a metal. The resin and Humic acid will compete for the metal and a certain proportion will bind to each (Baker, et al., 2007). By learning Schubert's method, we can find stability constants of complexes and coordination number through

* Corresponding Author
keilah2006@gmail.com

Advisor: Dr. Chulsung Kim
chulsungkim@gmail.com

mathematical methods. This ion exchange equilibrium method can almost be used to all the coordination reactions of Humic acid with any metal ions (Baker, et al., 2007; Qiang, et al., 2006).

Cation exchange equilibrium method is mainly based on the metal ion quantity (M_R), which adsorbed by a certain amount of cationic resin (R), is proportional to the metal ions in solution concentration (M) in a fairly large range of concentrations, when the reaction has reached equilibrium. In the systems, with or without a complexing agent, the metal ions distribution coefficient is λ_0 , λ respectively.

The distribution ratio, λ , of the metal ions between the cation exchange resin and the solution in the presence of a ligand is as follows:

$$\lambda = \frac{[M_R]}{[M]+[MA_x]} \quad \text{eq. (1)}$$

Where $[M_R]$ is the mole number of metal bound to per g of the resin (mol/g), $[M]$ is the concentration of free metal ion in the solution (mol/L) and $[MA_x]$ is the concentration of complexed metal (mol/L). $[M+M_x]$ is the total molar concentration of metal ions in the equilibrium solution (mol/L) (Baker, et al., 2007).

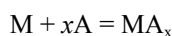
In the absence of ligand in solution, the distribution ratio, λ_0 , is given by:

$$\lambda = \frac{[M_R]}{[M]} \quad \text{eq. (2)}$$

The following equation can be derived:

$$\lambda = \frac{[M_R]}{\left[\left(\frac{1}{\lambda} - \frac{1}{\lambda_0}\right)\right]} \quad \text{eq. (3)}$$

Assume the equation of coordination reaction of Humic acid with metal ion is



M is the metal ion; A is the complexing agent; x is Coordination number. Stability constants of complexes K_f is:

$$K_f = \frac{[MA_x]}{[M][A]^x} \quad \text{eq. (4)}$$

Eq. (1) (2) (3) are reduced to:

$$K_f = \frac{[M_R] \left(\frac{1}{\lambda} - \frac{1}{\lambda_0}\right)}{\frac{[M_R][A]^x}{\lambda_0}} = \frac{\left(\frac{\lambda_0}{\lambda-1}\right)}{[A]^x} \quad \text{eq. (5)}$$

Logarithm on both sides, we have:

$$\log\left(\frac{\lambda_0}{\lambda-1}\right) = \log K_f + x \log[A] \quad \text{eq. (6)}$$

The pH, ionic strength, temperature, and cation exchange resin are under the condition of constant quality; changing the concentration of Humic acid (A_1, A_2, A_3, A_4), can be obtained with more corresponding values ($1, 2, 3, 4$). If the concentration of the complexing agent is much greater than the concentration of metal ion concentration, the initial concentration can be considered equal to the equilibrium concentration of free complexing agents. Thus, by using $\log\left(\frac{1}{\lambda} - \frac{1}{\lambda_0}\right)$ as the vertical axis; using $\log[A]$ as the abscissa map, stability constants of complexes ($\log K_f$) can be obtained from the intercept, and coordination number (x) calculated from the slope.

2. Reviews

2.1 The coordination reaction of Humic acid with Fe^{3+} ion

By doing experiments about effects of acidity on stability constants of complexes and the coordination number on coordination reaction of Humic acid with Fe^{3+} ion, the following linear relation was been obtained from Fig.1 and Fig.2 (Qiang, et al., 2006).

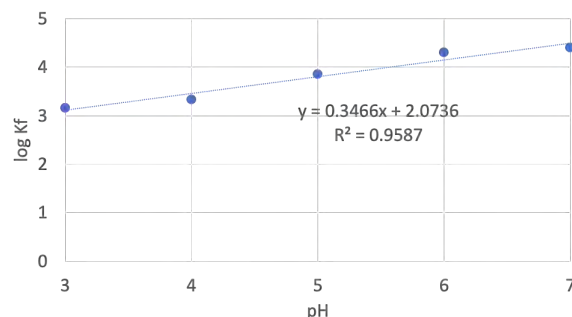


Fig.1 Effect of acidity on stability constants of complexes of HA - Fe^{3+}

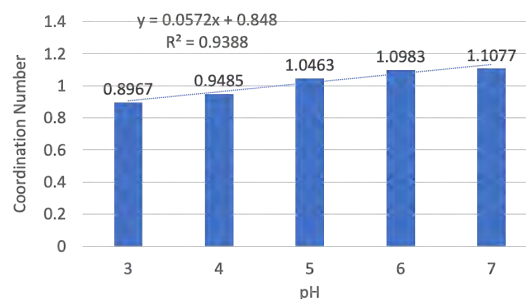


Fig.2 Effect of acidity on coordination number of HA- Fe^{3+}

The results show that the stability constants of complexes and coordination number are increasing as the pH is increased when ionic strength and temperature conditions are kept the same. The results suggest that a higher pH can increase the coordination reaction to the right.

Equations (7) and (8) are used to describe the effect of acidity on stability constants of complexes and coordination number of HA - Fe³⁺.

$$\log K_f = 0.3466 * \text{pH} + 2.0736 \quad \text{eq. (7)}$$

$$x = 0.0572 * \text{pH} + 0.848 \quad \text{eq. (8)}$$

The variation result from both acid effect of Humic acid and hydrolysis effect of Fe³⁺ ions. When the pH is low, the acid effect of Humic acid is very serious. It reduces the ability of the Humic acid ligand, resulting in a lesser extent of complex reactions to right, and resulting in small stability constants of complexes ($\log K_f$). With the gradual increase in pH, acid effect of Humic acid decreases; coordination ability gradually increases (Qiang, et al., 2006); the degree of coordination reactions to the right increases, so the stability constants of complexes ($\log K_f$) increase.

At the same time, with the hydrolysis effect of Fe³⁺ increasing, competitive conduct of OH⁻ will make the right degree of coordination reaction decreases, and thus the relative complexing ability of Humic acid declines, which leads to coordination number increases.

2.2 The coordination reaction of Humic acid with Mn²⁺ ion

By doing experiments about the effects of acidity on stability constants of complexes and coordination number on coordination reactions of Humic acid with the Mn²⁺ ion, from the experiment, Fig.3 and Fig.4 are obtained (Qiang, et al., 2008).

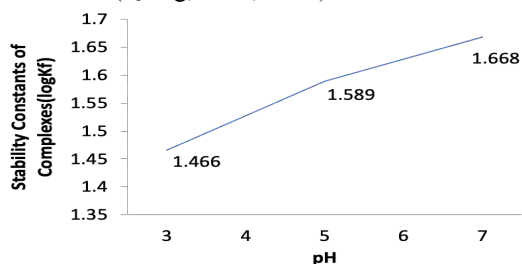


Fig.3 Effect of acidity on stability constants of complexes of HA - Mn²⁺

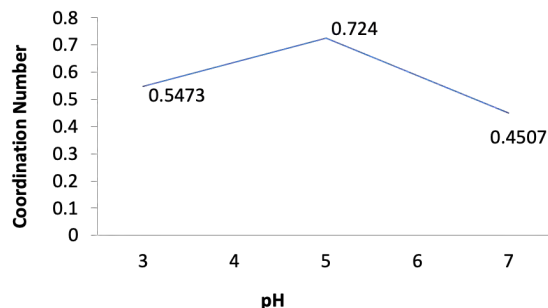


Fig.4 Effect of acidity on coordination number of HA- Mn²⁺

Effects of acidity to the coordination reaction of Humic acid with Mn²⁺ ion also can be analyzed by using the ion-exchange equilibrium method. The results show that the stability constants of complexes and coordination number are increasing as the pH is increased, if the ionic strength and temperature are under the condition of constant quality.

We can use equation (9) to describe the effect of acidity on stability constants of complexes ($\log K_f$) of HA - Mn²⁺.

$$\log K_f = 0.0504 * \text{pH} + 1.3225 \quad \text{eq. (9)}$$

Coordination number increased as pH is increased during the range of 3.0 to 5.0, but decreased when pH is increased from 5.0 to 7.0.

The variation result from both acid effect of Humic acid and hydrolysis effect of Mn²⁺ ions. When the pH is low, the acid effect of Humic acid is very serious. With the gradual increase in pH, it is the same influence to K_f just like the reaction of HA-Fe³⁺: acid effect of Humic acid decreases; the degree of complexation reactions to right increases. So the stability constants of complexes ($\log K_f$) increase.

At the same time, the hydrolysis effect of Mn²⁺ increases; the competition of OH⁻ increase the number of hydroxyl groups around the metal ions, changing the radius of hydration of metal ions, thus changing the coordination number of the coordination reaction (Qiang, et al., 2008).

2.3 The coordination reaction of humic acid with Cd²⁺ ion and Zn²⁺

By doing experiments about effect of acidity on stability constants of complexes and coordination

number on coordination reaction of Humic acid with Cd^{2+} and the Humic acid and Zn^{2+} respectively, following Fig.5 and Fig.6 and be drawn (Baofeng, et al., 2005).

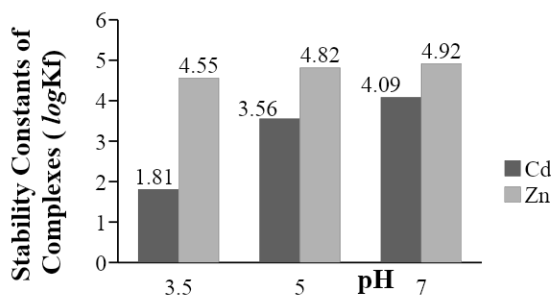


Fig.5 Effect of acidity on stability constants of complexes of HA-Cd^{2+} and HA-Zn^{2+} respective

Figure 5 shows, when pH increased from 3.5 to 7.0, stability constants of complexes ($\log K_f$) of HA-Cd^{2+} increased from 1.81 to 4.09, and stability constants of complexes ($\log K_f$) of HA-Zn^{2+} increased from 4.55 to 4.92.

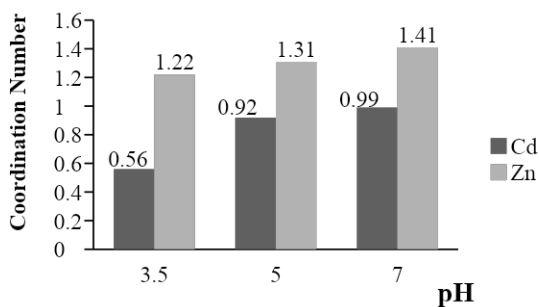


Fig.6 Effect of coordination number of HA-Cd^{2+} and HA-Zn^{2+} respective

Figure 6 shows, when pH increased from 3.5 to 7.0, coordination number of HA-Cd^{2+} increased from 0.56 to 0.99, coordination number of HA-Zn^{2+} increased from 1.22 to 1.41. In the two coordination reactions, both stability constants of complexes and the coordination number of HA-Cd^{2+} , and HA-Zn^{2+} will increase as the pH value is gradually increased.

This is because when the pH value is low, the concentration of H^+ in a solution is relatively high, certain weak binding sites in Humic acid are saturated, and thus the metal ions and Humic acid binding is inhibited. When pH value is increased, free functional groups, especially weak acidic carboxyl group, increases. The Humic acid molecule surface

electrostatic repulsion increases, therefore Humic acid changes its structure from reunion status to a stretched mesh structure when pH is low (Baofeng, 2005). Therefore metal ions are more accessible to the inner surface in order to combine with internal binding sites to enhance the stability of the reaction.

From Fig. 5 and Fig. 6, it is also concluded that both stability constants of complexes ($\log K_f$) and coordination number of HA-Zn^{2+} is greater than the constants of the reaction of HA-Cd^{2+} when they are in the same pH conditions. This is because the affinity ability of Humic acid and metal ions are related to the atomic coefficient and ionic radius. The ionic radius of smaller elements have a higher affinity with the HA. Element Zn has a smaller atomic coefficient and ionic radius than element Cd, so Zn^{2+} has a higher affinity with the HA than Cd^{2+} when they are in the same pH condition. Therefore, $\log K_{\text{Zn}}$ is larger than $\log K_{\text{Cd}}$.

3. Conclusions

Acidity affects both the acid effect of Humic acid and the hydrolysis effect of the Fe^{3+} ion. It creates opposite reaction directions, but not the same degree of influence. In certain experimental range, with the increase of pH, both stability constants of complexes and coordination number of HA-Fe^{3+} increases, which promotes coordination reactions. A higher pH value increases both stability constants of complexes and coordination number.

Acidity affects both the acid effect of Humic acid and the hydrolysis effect of the Mn^{2+} ion, just like the reaction of HA-Fe^{3+} . Higher pH value is positive to coordination reaction and increases stability constants of complexes. The coordination number increases with pH when it rises from 3.0 to 5.0, but decreases when the pH rises further up from 5.0 to 7.0.

To the reactions of HA-Zn^{2+} and HA-Cd^{2+} , both the stability constants of complexes and coordination numbers were increased as the pH is increased, when ionic strength and temperature conditions are kept the same. It is also concluded that both stability constants of complexes ($\log K_f$) and the coordination number of HA-Zn^{2+} is greater than the constants of the reaction of HA-Cd^{2+} when they are in the same pH conditions.

References

- Baker, H., & Khalili, F. (2007). Effects of pH and temperature on the interaction of Pb(II) with azraq humic acid studied with schubert's ion exchange method. *Annals of Environmental Science*, (1), 35-44
- Baofeng, L., Qiang, W., Wei, M., & Shiqiang, W. (2005). The impact of acidity on the complexation react of humic acid with Cd²⁺ and Zn²⁺. *Journal of Southwest Agricultural University (Natural Science)*, Aug. 27(4), 451-455
- Brigante, M., Zanini, G., & Avena, M. (2009). Effect of pH, anions and cations on the dissolution kinetics of humic acid particles. *Col. and Sur. a Physiochemical and Eng. Aspects*, 347, 180-186
- Dauletbay, A. et al. (2020). Interaction of metal ions with Humic acids of brown coals of Kazakhstan. *J. of Experimental Nanoscience*, 15, 406 - 416. <https://doi.org/10.1080/17458080.2020.1810240>
- Ghabbour, E.A., & Davies, G. (2001). *Humic Substances: Structures, Models and Functions*. Cambridge, U.K.: RSC publishing
- Qiang, W., & Shiqiang, W. (2006). The stability and thermodynamics of the coordination reaction of humic acid with Fe³⁺ ion. *Acta Scientiae Circumstantiae*, 26(1), 118-123
- Qiang, W., & Shiqiang, W. (2008). Characteristics of complex reaction between humic acid and Mn²⁺ ion. *Acta Pedologica Sinica*, 48(1), 47-53
- Stevenson, F.J. (1994). *Humus Chemistry: Genesis, Composition, Reactions*. John Wiley & Sons, New York
- Tipping, E. (1994). WHAM - a chemical equilibrium model and computer code for waters, sediments, and soils incorporating a discrete site/electrostatic model of ion-binding by humic substances. *Computers and Geosciences*, 20 (6), 973-1023
- Zhou, S. et al. (2015). Influence of Humic Acid Complexation with Metal Ions on Extracellular Electron Transfer Activity. *Sci. Rep* 5, 17067

On the Normality of the Distribution of Colors of m&m Candy

Ankita S. Achanta^{1*}

¹George School, Newtown, PA USA

Received August 31, 2022; Revised April 2, 2023; Accepted, May 6, 2023

Abstract

This investigation explores the shape of the distributions of colors of m&m candy. Statistical tests: kurtosis, skewness, and Shapiro-Wilk were used to assess normality of the distribution of the colors. Kurtosis measures “peakedness” and skewness measures distortion of a distribution. The Shapiro-Wilk test outputs a categorical yes/no, to indicate normality. Distributions of colors were also visually assessed for normality through histograms. This investigation also considers how normality varies based on sample size and different samples of the same size. It was hypothesized that if m&m candy colors are randomly allocated to packets, with an adequately large sample size, they would follow a normal distribution (skewness = 0, kurtosis = 3, Shapiro-Wilk test result = yes). The results support that increasing sample size does not make kurtosis and skewness closer to normally distributed with small samples ($n < 50$). Also, there is evidence that normality varies within samples of the same size, based on skewness/kurtosis. Finally, approximate probability density functions (PDFs) were created for each color distribution. With the functions, probabilities of getting a certain number of candies of a given color were calculated. Based on the PDFs, it was found that the probability of getting a packet with only blue m&ms is higher than that of a packet with only red m&ms, though both are highly unlikely. There are certain limitations related to this investigation. It is possible due to the small group sizes that the kurtosis/skewness values did not approach normality. Experimentation with larger sample sizes (group size > 100 each) would facilitate more accurate estimation of probabilities.

Keywords: Normality, Distribution, Kurtosis, Skewness, Shapiro-Wilk test

1. Introduction

M&ms are candy-coated chocolates distributed by Mars, Inc. They come in six colors: blue, orange, green, yellow, red, and brown. Usually, m&ms are distributed in small “fun-sized” packets with, on average, 15.9 candies per packet. This investigation explores the color distribution of the m&m candies. The intent is to evaluate if Mars, Inc. considers consumer color preferences and controls the distribution of each color to appeal to customers. Assuming no deliberate measures taken by Mars,

Inc., to control the colors of candy, the distributions should be random and thus normal. The null hypothesis is that Mars, Inc. did not intervene and the colors are normally distributed. The alternate hypothesis is that there was intervention and the distributions are not normal.

The primary aim of this study is to investigate the distribution of each of the six colors of m&ms. It employs statistical techniques to check normality and estimate the probability of obtaining any given number and color of m&m candies in a fun-sized packet.

* Corresponding Author
ankitaachanta@gmail.com

Advisor: Autumn Thayer
athayer@georgeschool.org

According to LaSalle University, m&m color blends were selected via consumer preference tests, which indicate the most attractive assortment of colors to people. Further, LaSalle reports that according to Mars, Inc., on average each packet of m&ms contains 24% blue, 13% brown, 16% green, 20% orange, 13% red, and 14% yellow m&m's.

1.1 Assumptions

Since Mars, Inc. produces all m&m packets, one wholesale bag was purchased instead of many small bags from different locations. It was assumed that the variability in the colors of m&ms purchased across wholesale bags or purchase points is trivial.

To calculate probabilities, a probability density function (PDF) for the distribution must be written. A PDF is “a function of a continuous random variable, whose integral across an interval gives the probability that the value of the variable lies within the same interval” (Oxford Languages, 2022). An assumption was made for the probability calculations. In order to approximate the probability of getting a single number calculated via integration, a correction is needed. For example, to calculate probability that proportion of blue m&ms is exactly 1, the boundaries of the integral were changed to (0.95,1.05).

Also, it was possible that the changes to kurtosis/skewness were not observable due to the small sample sizes of some groups.

2. Materials and Methods

A wholesale bag of fun-sized m&ms, labeled to contain 50 fun-sized packets and distributed by Mars, Inc. was used. An ideal statistical sample surveys 10% of the population as long as it does not exceed 1000 (MV Organizing 2021 para. 2). Since m&m distributes over 10 million packs per day, it would not be reasonable to sample 10% of all candies (about 1 million packs), so a sample size of 50 fun-sized bags was obtained for the study. While the label on the bag stated that there were 50 packets, the bag had 48 packets in reality, so a sample size of 48 packets was used. Microsoft (MS) Excel was used to create data tables and histograms with the frequency of each color in a packet.

The number of each color of m&ms in a fun-sized pack (blue, green, brown, orange, red, and yellow) was manually counted and recorded in MS Excel. Since the total number of candies per packet is variable, percentages (the percentage of the total number of candies in that packet that are the specific color of interest) were calculated to show frequencies of colors.

Normal distributions are perfectly symmetrical curves where the mean, median, and mode are equal. The total area under the curve of a standard normal distribution equals 1. Such distributions are important because they can be used to calculate the probability of a random variable taking a certain value or range of values. For example, the probability that a bag of fun-sized m&ms will contain 4 or more blue candies can be calculated using the probability density function (equation for the standard normal curve) for the random variable of the number of blue candies per packet. The probability density functions (PDF) were estimated for each color distribution to predict such probabilities.

Normality was tested in three different ways: kurtosis, skewness, and the Shapiro-Wilk test. An MS Excel add-in called Xrealstats (Real Statistics, 2022) was used to calculate these statistics and conduct the test.

“Kurtosis is a statistical measure that defines how heavily the tails of a distribution differ from the tails of a normal distribution” (Corporate Finance Institute,). In other words, kurtosis measures the “peakedness” of a distribution. The kurtosis of a normal distribution is always equal to 3.

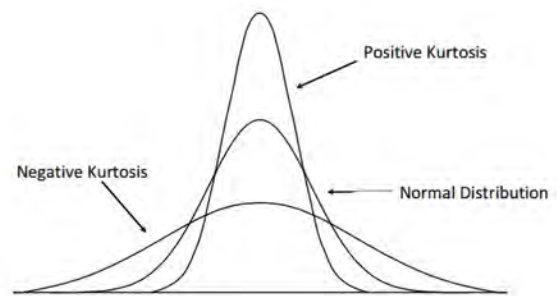


Figure 1: Examples of distributions with varying kurtosis (Glenn, 2015)

Positive kurtosis (leptokurtic) refers to a

distribution with a higher kurtosis than a normal distribution (kurtosis > 3). Negative kurtosis refers to a distribution with lower kurtosis than a normal distribution (kurtosis < 3). As shown in *Figure 1*, the distribution with a positive kurtosis is far more peaked than a normal distribution, and the normal distribution is more peaked than the one with a negative kurtosis. Also, data sets with high kurtosis have heavy tails, or outliers.

In MS Excel, the kurtosis formula is modified so that the kurtosis of a normal distribution is zero. To calculate kurtosis, standard deviation must be calculated first. Then, the following formula may be used:

$$\left\{ \frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum \left(\frac{x_j - \bar{x}}{s} \right)^4 \right\} - \frac{3(n-1)^3}{(n-2)(n-3)}$$

Note: n = number of terms in data set; x_i = observed data point; \bar{x} = sample mean; s = sample standard deviation

“Skewness refers to a distortion or asymmetry that deviates from the...normal distribution.” The skewness for a normal distribution is equal to zero. Negative values for skewness indicate data that are skewed left, and positive values indicate data that are skewed right (Investopedia, 2022).

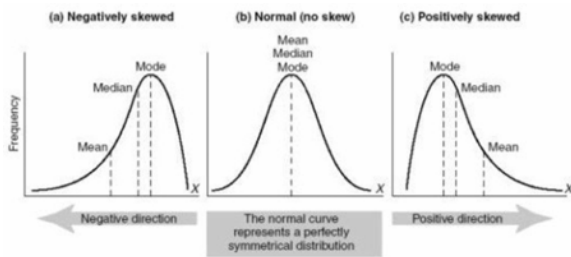


Figure 2: Examples of distributions with varying skews (Hogan, 2022)

Figure 2 shows how skewed distributions differ from normal distributions. Skewed data is not symmetric, and the mean, median, and mode are not equal. MS Excel’s skewness formula:

$$\frac{n}{(n-1)(n-2)} \sum \left(\frac{x_j - \bar{x}}{s} \right)^3$$

Note: n = number of terms in data set; x_i = observed data point; \bar{x} = sample mean; s = sample standard deviation

Finally, the Shapiro-Wilk test for normality outputs a categorical “yes/no” to indicate if the data is normal. The null hypothesis is that the variable is normally distributed, and the alternative hypothesis is that it is not normally distributed. The test produces a W statistic, along with a p-value. Testing at the 5% level of significance, if the p-value is less than 0.05, the null hypothesis is rejected and there is sufficient evidence to support that variable is not normally distributed (SPSS Tutorials, 2022).

These three tests are used to determine the normality of the groups of bags because they are readily available through Xrealstats.

For visual analysis, graphs were created with the data through MS Excel. Box plots were made for groups of bags. Also, histograms with normal curve overlays were created to check normality and write the approximate probability density function of each color.

The study investigated the following questions of interest:

Are the distributions of each color across packets normal?

Xrealstats was used to create a histogram with a normal curve overlay to visually assess normality, in addition to the kurtosis, skewness, and Shapiro-Wilk calculations. Further areas of inquiry that arise from this question are the following:

How does normality vary based on different samples of size 5 and 10 bags in each group?

Kurtosis, skewness, and Shapiro-Wilk test results of the first 10 groups of 5 bags for each color (bags 1-5, 6-10, etc., the last group was bags 44-48) will be compared to assess how normality varies based on different samples of the same size. The same tests will be used to compare variation in the normality of the 5 groups of 10 bags for each color.

How does normality vary based on the sample size (number of fun-sized bags)?

The first group of 5, 15, 25, 35, and 45 bags for each color will be taken from the data. The kurtosis, skewness, and Shapiro-Wilk test results for each color distribution will be compared to understand how normality changes.

Can the probabilities of obtaining certain numbers of certain colors be calculated with the estimated probability density functions?

The probability density functions of the color distributions will be estimated with sample data, and probabilities will be calculated using integration.

3. Results and Discussion

The raw data and percentages for the counts of m&m colors in each packet are available upon request from the first author. Further, kurtosis, skewness, box plots, and Shapiro-Wilk results for groups of 5/10 bags may also be requested.

3.1 Normality of the distribution of each color

Xrealstats was used to create the following histograms. The total area of the bars in the graph was used to create a normal curve overlay. Note that due to the bin sizing in the histograms below, the graphs show a bin less than zero; this is an artifact of the software used.

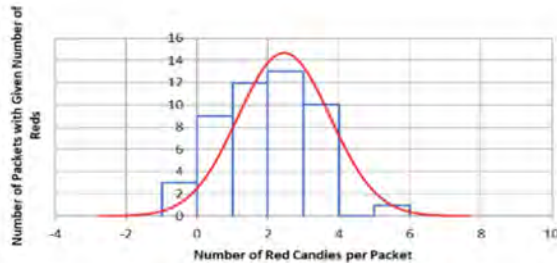


Figure 3: Red Color Distribution in 48 Packets

Visually, red is a near-normal distribution. The histogram is approximately symmetric, and the mode is around three red m&ms in a packet. If there were two or three packets with five red m&ms in them, the histogram would be even more symmetrical.

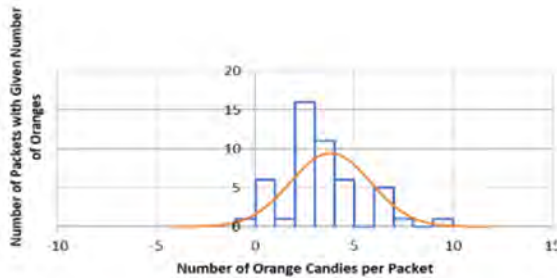


Figure 4: Orange Color Distribution in 48 Packets

The orange histogram looks slightly right skewed and is not as symmetrical as red. This is because the mode occurs at 2 orange candies per bag, and the distribution has larger values that are potential outliers (9 and 10). Also, the curve underestimates the mode of the histogram. The orange distribution is not normal.

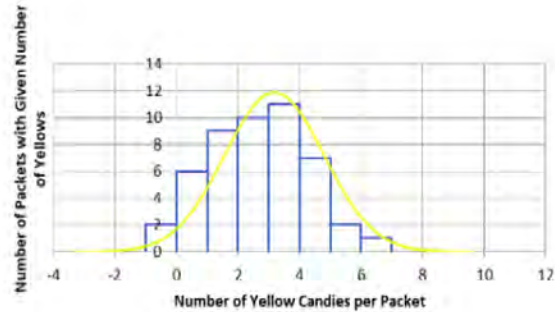


Figure 5: Yellow Color Distribution in 48 Packets

Yellow is the most near-normal distribution of the six colors. The curve seems to accurately represent the histogram, which is close to symmetrical and does not have any visible outliers. The mode is around three yellow candies in a packet.

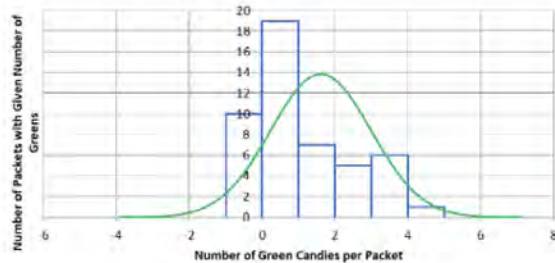


Figure 6: Green Color Distribution in 48 Packets

The green distribution is right-skewed. This may be because there were relatively fewer green candies per bag than the other colors, and with smaller sample size, the distribution showed a skew. The histogram shows the mode to be around 1, and the highest value of green m&ms per bag is 5, quite small compared to red, orange, and yellow. The curve does not seem to adequately represent the histogram, as the peak is far below the modal peak. Green is not normally distributed.

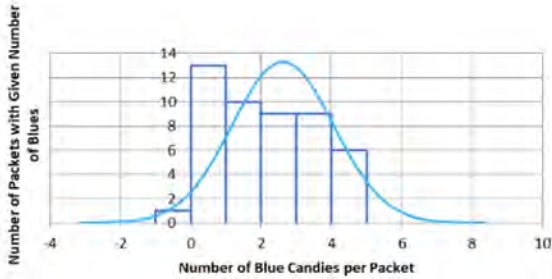


Figure 7: Blue Color Distribution in 48 Packets

The blue distribution has a mode around 1 and does not have more than 5 candies per bag. This histogram is uniform, but the curve does not represent the bars. The peak of the curve is around three, while the true mode is around one, so the blue distribution is not normal according to this data.

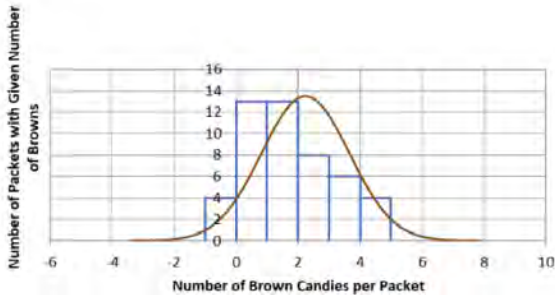


Figure 8: Brown Color Distribution in 48 Packets

The brown distribution is almost symmetrical. The modes are from 1-2, indicating that brown is a relatively less frequent color to occur in bags of m&ms. The curve almost matches the histogram, but because the bars are not perfectly symmetrical, brown is not normally distributed.

3.2 Variation in the normality of different samples of size 5 and 10 bags each

Observations from *Table 3* include that the kurtosis of red m&ms in groups of 5 bags varies from close to zero to almost three. This means that some of the groups of red are close to normally distributed (0) while others are slightly peaked (3). The 10th group of red bags has the most symmetrical box plot, but most of the plots are not approximately normal.

In orange, the kurtosis scores vary from -2.79 to 4.30, a wide range. Since many of the orange scores

were negative, the orange distribution likely has lighter tails than a normal distribution, meaning that most of the values are not extremely far from the mean. None of the orange box plots seem to be perfectly normal, and many are skewed or have outliers. These results may be discordant due to the small sample size of 5 bags (not large enough to be a representative sample).

The yellow kurtosis scores vary from -2.95 to 2.03 and are not near normal. Some of the box plots- 1,6, and 7 are symmetrical.

Green scores range from -3.04 to 2.06. Since there were very few green m&ms per packet, the green color distribution is expected to be the least normal of all of the colors due to the small sample size. This is reflected in the box plots as many are skewed or have wide ranges.

Blue kurtosis scores range from -2.69 to 1.21. There is one kurtosis score (0.16) that is close to normally distributed for the blue color. The box plots for the blue distributions have wide ranges and many are skewed.

Finally, the kurtosis scores for brown range from -3 to 4.7. Some distributions are very peaked for brown while others are not. Some of the box plots for the brown groups of 5 are symmetrical and look near normal. Kurtosis varied greatly among the distributions of colors for different groups of the same sample size. Also, skewness varied, but not as much as kurtosis. There was at least one group of bags with skewness of close to zero in each color, but there were more groups with skewness of around 1 overall.

For the distributions of groups of ten bags, the kurtosis varied less as the ranges were not as wide as the kurtosis of groups of 5 for each color. Across all colors, there were more negative values of kurtosis, so the distributions had lighter tails when the sample size was increased in each color. An increase in sample size decreases variability, so it makes sense that the distributions will have lighter tails with a larger sample size. However, the skewness seemed to be greater in groups of 10 bags, which was not expected. Overall, there were not as many values for skewness that were very close to zero. The box plots for each color were closer to symmetrical, especially the blue color distribution.

According to the Shapiro-Wilk test, the majority of the groups of 5 bags were normally distributed. However, results of this test were often discordant with those obtained using kurtosis and skewness. This may be due to the small sample sizes, and the experiment could be repeated with larger groups of bags for better results. This is a key consideration for analyzing the data as with smaller sample sizes, the Shapiro-Wilk test may not be accurate.

3.3 Variation in normality based on the sample size (number of fun-sized bags)

For the red distribution, the kurtosis gradually increases as the sample size increases. The skewness gets closer to zero, however, the group of five bags has a skewness of -0.02, which is closer to zero than the group of 45 bags, with a score of 0.06. Also, the Wilk-Shapiro test is not a good indicator of normality on its own, as all of the groups of bags were normally distributed according to the test.

The orange distribution's kurtosis gets closer to three (more peaked) as the sample size increases, with the exception of the first five bags as they start with a kurtosis score of close to three. Also, the skewness score gets closer to zero but does not decrease each time the sample size increases.

In the yellow distribution's results, kurtosis becomes closer to zero as sample size increases, getting closer to normal. Skewness is not very high and fluctuates slightly, it does not periodically decrease as sample size increases.

The green distribution's kurtosis scores get closer to normally distributed as the sample size increases. The skewness scores for green mostly approach zero as the sample size increases, and the closest score to zero is observed in the largest group of bags.

The blue kurtosis and skewness scores fluctuate as sample size increases, not showing a score that is close to normally distributed.

Finally, the brown group's kurtosis gets closer to normally distributed but is still far from the perfectly normal score of zero. The skewness for brown fluctuates.

3.4 Estimation of probabilities of obtaining certain numbers of certain colors candies

Estimated probability density functions for each color:

Probability Density Function (Duke Edu 2022):

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} * e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

μ and σ for populations are unknown, so they will be estimated as shown below:

$\sigma \cong s$ (sample standard deviation)

$\mu \cong \bar{x}$ (sample mean)

Estimated PDFs:

Red PDF: $\bar{x} = 0.153$; $s = 0.076$

$$F(x) = 5.001 * e^{-\frac{(x-0.153)^2}{0.0126}}$$

Orange PDF: $\bar{x} = 0.238$; $s = 0.122$

$$F(x) = 3.27 * e^{-\frac{(x-0.238)^2}{0.0297}}$$

Yellow PDF: $\bar{x} = 0.198$; $s = 0.099$

$$F(x) = 4.029 * e^{-\frac{(x-0.198)^2}{0.0196}}$$

Green PDF: $\bar{x} = 0.101$; $s = 0.0897$

$$F(x) = 4.447 * e^{-\frac{(x-0.101)^2}{0.016}}$$

Blue PDF: $\bar{x} = 0.165$; $s = 0.0911$

$$F(x) = 4.379 * e^{-\frac{(x-0.165)^2}{0.016}}$$

Brown PDF: $\bar{x} = 0.141$; $s = 0.092$

$$F(x) = 4.336 * e^{-\frac{(x-0.141)^2}{0.016}}$$

Using the PDFs above, some probabilities were calculated. First, the probability of getting one packet with all blue m&ms was calculated. Also, the probability of getting a packet full of red m&ms was calculated to be compared to the probability of getting all blues, since red candies were more frequent than blue ones.

Probability of getting a packet with all blue m&ms (Probability that proportion of blue m&ms in a randomly selected packet = 1)

$$\int_{0.95}^{1.05} 4.379 * e^{-\frac{(x-0.165)^2}{0.016}} dx = 8.27 * 10^{-19};$$

considered highly unlikely

Note: correction for the boundaries of the integral: (1-0.05), (1+0.05)

Probability of getting a packet with all red m&ms (Probability that proportion of red m&ms in a

randomly selected packet = 1)

$$\int_{0.95}^{1.05} 5.011 * e^{\frac{-(x-0.153)^2}{0.0126}} dx = 5.004 * 10^{-24}; \text{ also}$$

considered highly unlikely

It was found that the probability of getting a packet with all blue m&ms is relatively more likely than getting a packet of all red m&ms, however, both outcomes are considered to be highly unlikely.

4. Discussion & Conclusion

Based on the histograms, the closest to normally distributed color of m&ms is yellow, and red is also approximately normal. Further, at a constant sample size, kurtosis and skewness vary. Since kurtosis and skewness are random variables themselves, they have their own distributions. Further investigation focusing on kurtosis and skewness could examine if the distributions of kurtosis and skewness are the same as the distribution of the random variable used to calculate them.

Sample size does not make kurtosis and skewness closer to normally distributed based on these results. In some colors, there was an observed pattern between sample size and an increase in kurtosis, but many of the groups of bags were still far from normally distributed. It is possible that since the group sizes were small (5-45 bags), the changes to kurtosis and skewness were not observable. This experiment could be repeated with larger sample sizes to detect a more pronounced change in kurtosis by sample size.

The results demonstrate that the Shapiro-Wilk test lacks concordance with kurtosis and skewness scores. The kurtosis and skewness scores for many groups were not normally distributed, but the test statistic indicated that they were. Other groups that seemed to have kurtosis and skewness scores near normally distributed were not normal according to this test.

Due to the assumptions made and integral boundary corrections, this investigation has some limitations. Experimentation with larger sample sizes would be valuable in refining the probability density functions, thus calculating more accurate probabilities.

There may have been some intervention by Mars, Inc. in controlling the proportions of m&m colors as

many of the colors were not normally distributed. Thus, these findings support the alternate hypothesis.

References

(n.d.). Retrieved July 18, 2022, from <http://www.cs.uni.edu/~campbell/stat/prob9.html>

(n.d.). Retrieved July 18, 2022, from <https://sites.nicholas.duke.edu/statsreview/>.

Chen, J. (2022, July 8). *What is skewness in Statistics?* Retrieved July 18, 2022, from <https://www.investopedia.com/terms/s/skewness.asp#:~:text=Skewness%20refers%20to%20a%20distortion,is%20said%20to%20be%20skewed.>

Davis, B. (2021, April 30). *Home*. Retrieved July 18, 2022, from <https://www.mvorganizing.org/what-is-the-minimum-sample-size-for-experimental-design/#:~:text=1%2C200%20is%20desirable.>

Excel: How to calculate kurtosis with Kurt Function. (n.d.). Retrieved July 18, 2022, from https://indxar.com/contents/1404#:~:text=*The%20kurtosis%20of%20the%20normal,to%20be%200%20in%20Excel.

Glen, B. (2015) *Kurtosis Distributions*, Image, Stack Exchange,, <https://stats.stackexchange.com/questions/84158/how-is-the-kurtosis-of-a-distribution-related-to-the-geometry-of-the-density-function>

Hogan, R. (2020, June 16). *Probability distributions for measurement uncertainty*. isobudgets. Retrieved July 18, 2022, from <https://www.isobudgets.com/probability-distributions-for-measurement-uncertainty/>

Information Technology Laboratory. (2022, June 27). Retrieved July 18, 2022, from <https://www.itl.nist.gov/div898/handbook/eda/section3/eda35b.htm>.

Kurtosis. (2022, May 7). Retrieved July 18, 2022, from <https://corporatefinanceinstitute.com/resources/knowledge/other/kurtosis/>.

Measurement statistics: Color distribution in M&MS

- la salle university. (n.d.). Retrieved March 26, 2023, from <http://www1.lasalle.edu/~prushan/Measurement%20Statistics-lab2.pdf>

Microsoft. (n.d.). Retrieved July 18, 2022, from <https://support.microsoft.com/en-us/office/skew-function-bdf49d86-b1ef-4804-a046-28eaea69c9fa>
Oxford languages and google - english. Oxford

Languages. (n.d.). Retrieved July 18, 2022, from <https://languages.oup.com/google-dictionary-en/>.
SPSS Shapiro-Wilk Test – quick tutorial with example. SPSS tutorials SPSS ShapiroWilk Test Quick Tutorial with Example Comments. (n.d.). Retrieved July 18, 2022, from <https://www.spss-tutorials.com/spss-shapiro-wilk-test>

[-for-normality/#:~:text=The%20null%20hypothesis%20for%20the.sample%20from%20a%20normal%20distribution](#)

TodayShow. (2016, March 3). *Check out these 3 new peanut M&M's flavors, plus 5 M&M's fun facts*. TODAY.com. Retrieved July 18, 2022, from <https://www.today.com/food/5-things-you-didn-t-know-about-m-m-s-t77591>

Zach. (2021, March 2). *Can kurtosis be negative?* Statology. Retrieved July 18, 2022, from <https://www.statology.org/can-kurtosis-be-negative/>.

Zaiontz et al., Retrieved July 18, 2022, from <https://www.real-statistics.com/free-download/real-statistics-resource-pack/>.

Developments and Challenges with Earthquake Detection, Prediction and Protection

Jason Dong^{1*}

¹Troy High School, Fullerton, CA USA

Received October 14, 2022; Revised April 22, 2023; Accepted, May 25, 2023

Abstract

Earthquakes have influenced life on Earth since civilization began. With urbanization and increased population densities across the globe, earthquakes pose a high threat to life and infrastructure alike. Although as a society, we have not yet developed a consistent method of earthquake prediction, developing technologies that are fundamentally based on different fields should allow us to attack this problem from multiple fronts. Research into the rapidly growing field of geosciences, specifically geophysics, is more important now than ever. In this review, the history of humanity's documentation of earthquakes is explored, such as early earthquake quantification methods and seismograph development. Current detection methods and data analysis research into earthquake modeling are also reviewed here as they are currently our only means of prediction, and reveal the hopeful directions improved prediction in the future. Lastly, an interesting branch of earthquake prediction is explored, which involves animal behavior and the corresponding electric field phenomena correlated with earthquakes.

Keywords: Earthquakes, Earthquake Prediction, Earthquake Detection, Geophysics, Geosciences, Earth Science

1. Introduction

In the Chilean earthquake of 1960 that cost \$4.8 billion in damage and thousands of lives, seismic waves were recorded around the entire earth for multiple days (Encyclopedia Britannica). Each year 500,000 detectable earthquakes occur, 100 of them causing damage (USGS, Cool Earthquake Facts), and so it follows that religious, cultural, and political infrastructure developed around them, some of which can even be traced back to the ancient Egyptians. Once believed to be works of supernatural beings, earthquakes have become ingrained into human knowledge as both unpredictable and powerful. Although earthquakes have affected society since ancient times, effective detection and prediction of earthquakes remains shockingly minimal. This

review will discuss what is known about the science of earthquake detection and prediction, and explore current research efforts in this important field. Currently, large systems of seismometers provide large and connected databases to document earthquakes across the globe. This data then is used to fuel research in unearthing predictive signals- such as recognizing changes in the atmosphere or unearthing complex patterns in ambient seismic fields. These research efforts are unfortunately, still fairly incomplete, therefore the limitations and challenges of this important field will be explored below.

Many historic interpretations of earthquakes were based on religion, the first modern scientific conjecture of the cause of earthquakes was made in 1760 by John Mitchell, who proposed that "shifting

* Corresponding Author
jasondong.ai@gmail.com

Advisor: Dr. Heather Killeen
am.HeatherKilleen@gmail.com

masses of rock miles below the surface” release the waves of motion that are understood as earthquakes, which ripple across the conductive surface of the Earth (USGS, Cool Earthquake Facts). As it is understood today, earthquakes occur when a build-up of seismic stress along tectonic plates of rock release, causing the blocks to violently move relative to each other.

Current solutions to reduce damage of earthquakes include developing resistant architecture that can withstand the shaking force during an earthquake strike. There has been evidence of earthquake-resistant structures in what is now Iran from around 600 BCE, which demonstrate the isolation of different floors to help reduce damage when earthquakes hit (Patowary, 2019), and structures in earthquake-active Peru utilized dry-stone walls to disperse vibrations that might prove otherwise harmful to buildings at around the 13th century (Patowary, 2019). Earthquake engineering however, did not begin as a branch of Civil Engineering until the second half of the twentieth century (Reitherman, 2012). This period of time saw a number of catastrophic earthquakes, including the notable 1906 earthquake in San Francisco California (History.com, 2009). With modern times came more complex problems that drove engineering to minimize risks of downed wires, cracked gas lines, and broken stoves that have had a visible effect on appliances and architecture today.

The general public has also been trained on how to respond in case an earthquake does happen, but the truth is, humans still have very little power over an earthquake’s effects. What makes earthquakes so devastating is that earthquake prediction, the ability to tell us when one will hit and how strong it will be, is still a long way from being achieved. Therefore, traditional methods of evacuation fail in this type of disaster, which plays a significant role in the loss of life it brings. Until earthquakes can be accurately predicted, their dangers will remain as a considerable threat to society therefore, it is important for the current challenges and progress to be carefully explored.

2. History of Earthquake Documentation

Our grapple against the ground began with the oldest written records of earthquakes; these can be traced back to over three thousand years ago (USGS, Earthquakes in History). While man has likely always wondered why earthquakes occurred, notable progress was made in 132 BCE China with the invention of the first seismograph (Figure 1)(Kids Discover, 2014). The design was a vase with eight dragon heads in a circle that was sensitive to tremor direction by its eight sides (Figure 1). When slight tremors passed by, the device would drop balls in different directions (Figure 1). This device was a crucial step on the offensive: it was the first recorded mechanism to record the direction and magnitude of an earthquake.



Figure 1. A photograph of one of the first documented seismometers from China (Photo: Houfeng Didong) (Gillan, 2019).

Further progress would be made in terms of analysis, but the semblance of modern measures would only begin in 1889. In that year, German astronomer Ernst von Reber-Paschitz, while measuring gravitational attraction caused by other planets with a large pendulum, noticed a seemingly spontaneous and regular swinging unaccounted for by the planets (Berkeley, 2010). The cause of this force was unknown until he had read a report on a massive earthquake months later. Although his pendulum was based in Germany, it had recorded an earthquake all the way from Japan (Berkeley, 2010). Thus, Reber-Paschitz had accidentally laid the foundation for seismograms for centuries to come. Today, a three-story tall spring-pendulum design can

still be found functioning in Mexico City (Figure 2). Ultimately, research towards future predictions of earthquakes have all utilized data generated by seismograms, making this tool fundamental to making progress towards enhanced prediction.



Figure 2. A spring pendulum that is suspended three stories above from the roof of the Cathedral Metropolitana in Mexico City (Photo Credit: Diego Delso, License: <https://creativecommons.org/licenses/by-sa/3.0/>) (Wikiwand).

The 1906 San Francisco earthquake was a turning point for driving the scientific study of the San Andreas fault system in California. While significantly behind earthquake research from Japan and Europe, the first U.S. seismographs were installed in 1887 (USGS. 1906 Marked the dawn of the Scientific Revolution). Catastrophic earthquakes ultimately led to the first government-funded scientific investigation into earthquakes, the State Earthquake Investigation Commission, which published a report in 1908 that compiled knowledge from over twenty researchers and serves as a basis for what is now known about earthquakes. To this day, the Lawson report is still held in high regard.

3. Detecting Earthquakes Today

While knowledge of earthquakes has grown a lot since the first seismometer was invented, the basic tools used to detect earthquakes haven't changed a great deal. Seismometers today still utilize Ernst's principle of oscillating pendulums sensitive to ground movement. There are differences, though; since now

it is understood that earthquakes consist of four waves with different directions of movement, modern seismometers now include three axes of measurement, modeled by three different pendulums. A Z-component measures vertical motion, while two E/N components measure northern, eastern, western, and southern movement (British Geological Survey).

Another advancement is the presence of easily transportable seismometers today (Figure 3) that can be found at different locations across the globe help pinpoint the epicenters of earthquakes. Reduced cost and internet access have also given rise to many "public seismograph networks", one of which is funded by the U.S. government – the United States Geological Survey (USGS). Larger arrays have also been developed to map the interior of the Earth in terms of capturing wave speed and transmissivity.



Figure 3. Modern seismometer used at the USGS (Public Domain) (USGS).

There are thousands of seismometers now scattered throughout the world creating a network of databases of seismic data. The USGS is just one organization that cooperates with international partners such as the National Science Foundation, to generate an expansive public database of seismic events. Known as the Incorporated Research Institutions for Seismology (IRIS), this database offers data from even the early 1900s (USGS. Earthquakes in History). The ubiquitous nature of the internet also gilds these databases with an important function in early warning systems, where monitoring systems can immediately send electronic signals to damage-prone areas minutes before damage may be

done (Japan Meteorological Agency). This is based off of the principle that seismograms gather two types of wave data. The first is the longitudinal and comparatively harmless P-wave, or pressure wave, and the second is the more dangerous S-wave, or shear wave, which measures transverse movement. The p-wave of an earthquake travels about 67% faster than the dangerous S-wave, therefore the p-wave can be used to estimate the center of an impending shock, and areas that will be hit can be notified instantaneously (USGS, *Body Waves Inside the Earth*). Countries such as Japan have used these signals to broadcast expected earthquakes through TV and radio (Center for Public Impact). While this has made some level of evacuation possible, it would be valuable to increase the evacuation time to provide more effective evacuation of heavily populated areas and of citizens that require more time, leaving earthquakes as a substantial threat today.

4. Modeling Future Earthquakes

Global databases and systems of seismograms have allowed researchers internationally to conduct advanced studies of earthquakes in order to uncover novel methods to predict them. One particular area of research involves analysis of patterns in Earth's tremors has led to fascinating new insights. One study from Stanford examined how the ambient seismic field could hint at moderate-level earthquakes (Denolle et al., 2013). This study, as well as others (Ma et al., 2008) (Prieto et al., 2011), extracts impulse response functions over time between two measuring stations, then modifies and solves this system of differential equations to get useful information relating ambient seismic field patterns and larger seismic events. This is part of a larger effort to analyze earthquakes; however, it has been concluded that additional sensitive seismograph data would be needed to generate more accurate, and useful, functions. In another example, a group from UCLA analyzed strain data to determine appropriate models for earthquake impacts (Snieder 2006). More recently, an effort by the Bulletin of the Seismology Society of America compared station-to-station velocity models to actual data to analyze flaws in Green analysis (Paul et al., 2020).

Another approach that characterizes earthquakes is the analysis of the earthquake aftershock. Such is the tendency of a large earthquake to induce a series of smaller earthquakes on the same fault, an effect of the "readjustment process" of tectonic plates (USGS. What is the difference between Aftershocks and Swarms?). The Epidemic Type Aftershock Sequence (ETAS) model proposed by Ogata in 1988 (Ogata, 1988) is a popular stochastic model for seismicity, but despite decades of research this model lacked the ability to link individual seismic events. Therefore an alternative approach has arisen, known as the "earthquake productivity law", which is defined as the exponential distribution of related earthquakes independent of magnitude (Baranov, 2019). To measure "earthquake productivity" data trees are generated for related earthquakes and the number of "neighbors", or related events, for each earthquake, is graphed (Shebalin, 2020). An alternative model for earthquake aftershocks based on this has been developed and verified in light of other laws, and potential methods to estimate the magnitude of the largest aftershock of an event have been reviewed (Baranov, 2022).

Although such research has vastly expanded our understanding of earthquakes, more work needs to be done to leverage that understanding into useful prediction models.

5. Can Animals Predict Earthquakes?

Studies on earthquakes today encompass a far broader range of symptoms that might predict an earthquake and use increasingly novel strategies to do so. In a 2014 study by Hiroyuki Yamauchi, surveys revealed that pets were more restless before the earthquake prior to its occurrence (Yamauchi et al., 2014).

Records of unusual wildlife activity before earthquakes can be traced back to Greece in 373 BC (USGS, *Can animals predict earthquakes?*), in which centipedes and rats were reported to have left the city for safety. While people have noticed unusual animal behavior before earthquakes in recorded accounts (including a zoo in China that observed zebras banging their heads against walls and elephants wildly swinging their trunks before a large

earthquake) little has been done to quantify this phenomenon until recently (Sanderson, 2008). Surveys by Hiroyuki Yamauchi, Hidehiko Uchiyama, Nobuyo Ohtani, and Mitsuaki Ohta published in 2014 found a relationship between both distance and time from the epicenter with pet activity, reporting that “the ratio of total unusual animal behaviors increased in both [cats and dogs]” before a magnitude 9.0 earthquake occurred (Yamauchi et al., 2014).



Figure 4. Frog migrations before an earthquake in China (Mui, 2008).

Whilst the exact reason animals can detect this specific danger is unknown, current hypotheses include changes in “atmospheric pressure, gravity, ground deformation, noise from the formation of microcracks, groundwater level changes, and the emission of gases and chemical substances (Mui, 2008).” A 2017 study confirms the ability of animals to pick up on a large range of abiotic noises (Garstang, 2017). However, it contends that the sounds heard by animals before an earthquake are the result of induced vibrations in metal, the product of a large chain of activity starting with crustal breakage, not sounds directly from the crust breaking (Garstang, 2017). Such studies provide insight into unexpected signals from earthquakes- that alternative signals may be found, not necessarily in seismic vibrations. China, inspired by these findings, launched its own animal monitoring program in 2015 (BBC News, 2015). While no earthquakes have yet been predicted with this program, active analysis of large samples of animals may perhaps save lives in the future.

These efforts prompt the question: What are animals sensing? As seen, one of the hypotheses

includes low-frequency sounds outside of human hearing ranges. The answer may be in the sky rather than the ground below. Many studies have noticed correlative observations between earthquakes and electromagnetic wave emission, electric field phenomena, and magnetic field phenomena. A 2018 paper by Friedemann Freund, a NASA researcher, proposes a unifying theory, suggesting that tectonic stresses cause “the activation of electronic charges ... via the rupture of peroxy bonds (Freund et al., 2021)”. These electronic charges immediately travel fast and far out of stressed zones to trigger a handful of phenomena, including but not limited to those mentioned before, massive air ionization, and increased levels of ozone and carbon monoxide (Freund et al., 2021).

As for animals, Freund says that ionization of air molecules- which may, for example, oxidize water into hydrogen peroxide- as well as the oxidation of organic compounds, may cause unusual reactions within animals (Freund & Stolc, 2013). Whatever the pathway animals use to detect earthquakes, there is clearly more for us to learn from them to understand the subtle symptoms that precede an earthquake.

6. Conclusion

The 1906 San Francisco earthquake shook down \$15 billion dollars in infrastructure, ending 3000 lives and leaving 250,000 homeless (History.com, 2009). Models by USGS suggest that such catastrophic earthquakes occur every 200 years in the California region, therefore it is highly likely that California will experience one in the next few decades (USGS, When will it happen again?). Although countries like Japan currently use earthquake data, and specifically p-wave data, for prediction, the amount of advanced warning time could be vastly improved. Likewise, advances in ambient seismic data analysis have revealed valuable insights into moderately-sized earthquakes, however refinement of our current tools are needed for more sensitive measurements to facilitate prediction of larger quakes. New approaches to geophysical research are sure to lead to exciting new models of earthquake prediction, and one of these approaches may be through cluster analysis. For example,

tree-based analysis of aftershocks has already led to new insights into the nature of earthquakes, so as mathematical and computational science continues to evolve, we will develop stronger tools for more sophisticated analysis of seismic data. Lastly, to the more we learn about changes in animal behavior preceding earthquakes, the clearer the mechanisms will be and more thoroughly explored for further insight, leading us closer to improved earthquake prediction. With urbanization and increased population densities across the globe, posing a high threat to life and infrastructure alike, earthquakes have grown as influential as ever, making these developments all the more important.

Acknowledgment

I would like to thank Heather Killeen, Ph.D., and Justin Senter for their assistance in preparing this manuscript for publication.

References

- Baranov, S. V., *et al.*, (2019). Epidemic Type Aftershock Sequence Exponential Productivity. *Russian Journal of Earth Sciences*. Vol. 19, ES6003.
- Baranov, S. V., Narteau, C., Shebalin, P. (2022). Modeling and Prediction of Aftershock Activity. *Surveys in Geophysics*. Vol. 43, pp. 437-481. doi.org/10.1007/s10712-022-09698-0.
- BBC News (2015). China: Animal Centres Used to Help Predict Earthquakes. *BBC News, By News From Elsewhere*.
- Berkeley, UC. (2010). Today in Earthquake History: Japan/Potsdam 1889. *UC Berkeley, Berkeley Seismology*. seismo.berkeley.edu/blog/2010/04/17/today-in-earthquake-history-japan-potsdam-1889.html
- British Geological Survey. How are earthquakes detected? (accessed 2022 -04 -05).
- Center for Public Impact, J-Alert: disaster warning technology in Japan. (accessed 2022 -04 -05).
- Denolle, M. A., *et al.* (2013). Ground Motion Prediction of Realistic Earthquake Sources Using the Ambient Seismic Field. *Journal of Geophysical Research: Solid Earth* 2013, 118 (5), 2102–2118. doi.org/10.1029/2012jb009603.
- Freund, F.; *et al.* (2021). Earthquake Precursors in the Light of Peroxy Defects Theory: Critical Review of Systematic Observations. *The European Physical Journal Special Topics* 2021, 230 (1), 7–46. doi.org/10.1140/epjst/e2020-000243-x.
- Freund, F.; Stolz, V. (2013). Nature of Pre-Earthquake Phenomena and Their Effects on Living Organisms. *Animals* 2013, 3 (2), 513–531. doi.org/10.3390/ani3020513.
- Garstang, M.; Kelley, M. (2017). Understanding Animal Detection of Precursor Earthquake Sounds. *Animals* 2017, 7 (12), 66. doi.org/10.3390/ani7090066.
- Gillan, J. (2019). Ancient Chinese Earthquake Detector Invented 2,000 Years Ago Really Worked! *Ancient Origins*.
- History.com Editors. (2009). San Francisco Earthquake of 1906. *HISTORY*. Encyclopedia Britannica. Chile Earthquake of 1960. *Encyclopedia Britannica*.
- Japan Meteorological Agency. <https://www.jma.go.jp/jma/en/Activities/eew.html> (accessed 2022 -04 -05).
- Kids Discover. (2014). Ancient Chinese Seismometer Used Dragons and Toads. *Kids Discover*.
- Ma, S.; Prieto, G. A.; Beroza, G. C. (2008). Testing Community Velocity Models for Southern California Using the Ambient Seismic Field. *Bulletin of the Seismological Society of America* 2008, 98 (6), 2694–2714. doi.org/10.1785/0120080947.
- McAdam, D. (2019). New Technology Could Help Scientists Predict Earthquakes. *NBC San Diego*.
- Mui, M.; (2008). Frog Migration: Omen to China Earthquake Disaster. *Weird Asia News*.
- Ogata, Y. (1988). Statistical Models for Earthquake Occurrences and Residual Analysis for Point Processes. *Journal of American Statistical Association*. Vol. 83, No. 401, pp. 9-27.

Paul, C.; *et al.* (2020). Testing an Empirical Green's Function Method for Determining the Rupture Parameters of the 24 April 2014 Vancouver Island Earthquake. *Bulletin of the Seismological Society of America* 2020, 111 (2), 1043–1055. doi.org/10.1785/0120200233.

Patowary, K. (2019). Tomb of Cyrus: The World's Oldest Earthquake Resistant Structure. *Blogger*. World History Encyclopedia. Inca Architecture. *World History Encyclopedia*.

Prieto, G. A. *et al.* (2011). On Amplitude Information Carried by the Ambient Seismic Field. *Comptes Rendus Geoscience* 2011, 343 (8–9), 600–614. doi.org/10.1016/j.crte.2011.03.006.

Reitherman, R. (2012). Five Major Themes In the History of Earthquake Engineering. *Consortium of Universities for Research in Earthquake Engineering*.

Sanderson, B. H. (2008). Did Animals Predict China's Earthquake? *NBC News*.

Shebilin, P. N., Narteau, C., Baranov, S. V., (2020). Earthquake Productivity Law. *Geophysical Journal International*. Vol 222, Issue 2, pp 1264-1269. doi.org/10.1093/gji/ggaa252.

Snieder, R. (2006). Extracting the Building Response Using Seismic Interferometry: Theory and Application to the Millikan Library in Pasadena, California. *Bulletin of the Seismological Society of America* 2006, 96 (2), 586–598. doi.org/10.1785/0120050109.

USGS. 1906 Marked the Dawn of the Scientific Revolution. earthquake.usgs.gov/earthquakes/events/1906calif/18april/revolution.php (accessed 2022 -04 -05).

USGS. Body waves inside the Earth. earthquake.usgs.gov/earthquakes/events/1906calif/18april/earthwaves.php#:~:text=Two%20kinds%20of%20waves%20are,to%20arrive%20from%20the%20earthquake (accessed 2023 -04 -15)

USGS. Can animals predict earthquakes? www.usgs.gov/faqs/can-animals-predict-earthquakes? (accessed 2022 -04 -05).

USGS. Cool Earthquake Facts. www.usgs.gov/programs/earthquake-hazards/cool-earthquake-facts

USGS. Earthquakes in History. pubs.usgs.gov/gip/earthq1/history.html (accessed 2022 -04 -05).

USGS. International Programs. www.usgs.gov/international-programs (accessed 2022 -04 -05).

USGS. Seismometer. www.usgs.gov/media/images/usgs-seismometer (accessed 2022 -04 -05).

USGS. What is the difference between Aftershocks and Swarms? www.usgs.gov/faqs/what-difference-between-aftershocks-and-swarms#:~:text=Aftershocks%20become%20less%20frequent%20with,earthquakes%20with%20no%20identifiable%20mainshock. (accessed 2023 -04 -15).

USGS. When will it happen again? earthquake.usgs.gov/earthquakes/events/1906calif/18april/whenagain.php (accessed 2022 -04 -05).

Wikiwand. Pendulum. *Wikiwand*.

WorldNuclear.org, Fukushima Daiichi Accident, 2022. (accessed 2022 -04 -05).

Yamauchi, H.; *et al.* (2011). Unusual Animal Behavior Preceding the 2011 Earthquake off the Pacific Coast of Tohoku, Japan: A Way to Predict the Approach of Large Earthquakes. *Animals* 2014, 4 (2), 131–145. doi.org/10.3390/ani4020131.

Elucidating the Epigenetic Mechanisms that Modulate Biological Age

Salem Wang^{1*} and Arianna Broad²

¹North Farmington High School, Farmington Hills, MI, USA

²Weill Institute for Cell and Molecular Biology, Cornell University, Ithaca, NY, USA

Received November 6, 2022; Revised May 7, 2023; Accepted, June 2, 2023

Abstract

In recent years, changes in the epigenome have emerged as a key component in the aging process. DNA methylation, histone acetylation and methylation, and chromatin landscape remodeling are the most well researched of the types of epigenetic modifications. To best elucidate their functions, they should be understood as interacting parts in the context of an integrated epigenetic approach and not as exclusive mechanisms. Each type of modification undergoes some predictable changes with aging but may display variation depending on the individual, cell type, or organism. These epigenetic modifications correlate to changes in phenotypes common with aging and to increased susceptibility to age-related diseases, such as cancer or Alzheimer's Disease. While age may cause these changes, epigenetic modifications also contribute to aging by altering gene expression and interacting with other agents of aging to further the progression. Therefore, the scope of this review looks at epigenetic modifications as an effect and driver of aging.

Keywords: Aging, Epigenetics, DNA Methylation, Histone Modification, Chromatin, Chromatin Remodeling, Chromatin Landscape

1. Introduction

Cellular and molecular damage to cells over time accumulates in a phenomenon called aging, which is partly dependent on the passage of time, or chronological age. Biological age differs from chronological age, as it is the age of a person demonstrated by cellular health and is a more accurate predictor of lifespan and healthspan, the amount of time an individual lives healthily. The rate of aging was previously believed to be predetermined by the genes inherited at birth, but recent research has shown that the environment and the resulting epigenetic changes to the chromosomal landscape determine a significant portion of longevity and health (Kirkwood, 2005; Sen, et al., 2016). Aging is

also accompanied by modification of an organism's phenotype, which is connected to the epigenetic modulation of transcription (Kirkwood, 2005; Johansson, et al., 2013). Various other hypotheses and hallmarks of aging have also been proposed, including genomic instability, telomere attrition, and cell senescence (López-Otín, et al., 2013). These hallmarks of aging should not be seen as mutually exclusive, but as an interconnected network of processes influencing each other and contributing to aging as a whole. However, this review will concentrate mainly on epigenetic modifications as the most crucial factor in aging because it controls gene regulation and phenotypic alterations, thus making it an underlying factor that many of the other hallmarks of aging possess.

* Corresponding Author
wangsalem.m@gmail.com

Advisor: Arianna Broad
amb647@cornell.edu

Epigenetics refers to changes made to the chromosomal landscape without altering the genetic code directly and includes DNA methylation, histone modification, and chromatin remodeling. While there are other factors that influence epigenetic modification, such as modification of non-coding RNAs, which has also been established as an epigenetic factor more recently, this review will focus on the more established canonical methods of epigenetic modification. The most well studied epigenetic modification is DNA methylation, which refers to the addition of methyl groups to the cytosine in CpG dinucleotides, where a cytosine precedes a guanine in the 5' to 3' direction, and generally represses gene expression through decreasing access to gene regulatory elements, namely promoters and enhancers (Bormann, et al., 2016). Hypomethylation leads to reduced genomic stability and excessive gene expression, increasing the rate of mutations and the probability of age-related health issues, including cancer. Conversely, hypermethylation represses gene expression and may occur at protein-coding genes that carry out basic cell functions, henceforth depleting the cell of said protein which could have catastrophic consequences (Ashapkin, et al., 2017). Histone modification, defined as a post-translational modification (PTM) of histone proteins in chromatin, while less researched than DNA methylation, has also emerged as an important epigenetic factor in the remodeling of the genomic landscape. Amongst the various types of histone modifications, histone acetylation and histone methylation in particular have shown consistent correlations with aging. Both DNA methylation and histone modification are inextricably tied to chromatin remodeling, since DNA methylation correlates with decreased chromatin accessibility while histone modifications can dictate chromatin structure (Johansson, et al., 2013; Kouzarides, 2007).

As the average lifespan increases in developed countries, health in old age becomes an increasingly important concern. While it is not clear the order in which aging hallmarks arise, epigenetic research may be a crucial piece in improving healthspan in older age. The epigenome is alterable unlike the genetic code; therefore, it is theoretically possible to reverse some age-related epigenetic changes to revert a cell

to a biologically younger age. Pinpointing the exact epigenetic modifications that affect biological age may aid in the development of pharmaceuticals that target age-related epigenetic modifications to improve healthspan.

Many of the studies on epigenetics and its role in aging have been done in model organisms such as budding yeast (*Saccharomyces cerevisiae*) and the fruit fly (*Drosophila melanogaster*), which have short generations and high reproduction rates, providing ideal conditions for researchers to track target genes and phenotypes related to aging over many generations. These organisms also possess some highly conserved cellular processes, allowing us to make inferences about aging in humans (Kaeberlein, et al., 2007). To elucidate the importance of epigenetic mechanisms in the process of aging, this review highlights patterns of DNA methylation, histone modification, and chromatin remodeling as it relates to other hallmarks of aging, providing a holistic perspective on epigenetic modifications and their roles in aging. This holistic review will bring to light the extensive interplay between different epigenetic modifications, allowing scientists to easily identify the key questions in the field and investigate them experimentally.

2. DNA Methylation

As aging progresses, a global pattern of hypomethylation occurs, but specific CpG sites show hypermethylation (Day, et al., 2013). Aging is accompanied by an increase in heterogeneity between methylomes and amongst cells of the same tissue over time due to random errors in DNA methylation that may arise during replication (Winnefeld and Lyko., 2012; Bormann, et al., 2016; Ashapkin, et al., 2017). The increasing differences between individual methylomes account for many of the phenotypic differences that arise in identical twins as they progress through life, and the increasing heterogeneity within tissues increases risk of organ failure (Tan, et al., 2016; Ashapkin, et al., 2017). Studies in *D. melanogaster* and *C. elegans* have contributed to our understanding of the correlation of DNA methylation and aging, however, the type of DNA methylation that occurs in *D. melanogaster* and

C. elegans differs from the traditional mammalian (5mC) methylation (Booth and Brunet, 2016; Greer and Shi, 2012). Therefore, while model organisms can provide insight into the global correlations of DNA methylation and aging, DNA methylation must be studied as well in mammalian cells to provide insight into the cellular mechanisms behind epigenetic patterns and aging.

Tracking DNA methylation at CpG sites has been established as a reliable way to measure chronological age, more so than other well studied age predictors such as telomere length (Hannum, et al., 2012; Horvath, 2013; Jylhävä, et al., 2017). Although there are individual variations in patterns of DNA methylation, some methylation sites are highly conserved among mammalian species, demonstrating a clear link between aging and specific genes (Li, et al., 2022). The *ELOVL2* gene, which is associated with the synthesis of long fatty acid chains primarily in the liver, has been identified as the strongest correlative gene between methylation and age (Li, et al., 2022; Spólnicka, et al., 2018; Garagnani, et al., 2012). As chronological age increases, the *ELOVL2* gene becomes hypermethylated, decreasing gene expression (Garagnani, et al., 2012). Other CpG sites with methylation patterns highly correlated to age have been found to play a role in cancer and Alzheimer's Disease, among other common symptoms of aging and decline in health (Spólnicka, et al., 2018). Some CpG sites have shown strong positive correlation between DNA methylation and age in certain regions of the brain (Hernandez, et al., 2011).

When methyl groups bind to CpG dinucleotides, the new methylated form of cytosine called 5-methylcytosine (5mC) (Figure 1) inhibits binding of transcriptional activators or promotes binding of transcriptional repressors to repress gene expression (Watt and Molloy, 1988; Booth and Brunet, 2016). CpG sites are often located at the promoter regions of housekeeping genes, therefore regulating gene expression of essential proteins (Day, et al., 2013). However, methylation correlated with aging most dramatically occurs not in the promoter regions, but in the enhancer regions, which are also sites important to regulating gene expression (Johansson, et al., 2013). Furthermore, methylation at noncoding

sites has emerged in recent years as a significant factor in genome stability (Winnefeld and Lyko, 2012). It is thought that hypomethylation of noncoding DNA sites decreases chromatin density and allows insertion of transposable elements, thereby increasing genomic instability and mutations prevalence (Pal and Tyler, 2016). Therefore, breast and other types of tissue that typically indicate a higher biological age according to DNA methylation patterns also demonstrate a higher incidence of cancer and tumors (Horvath, 2013).

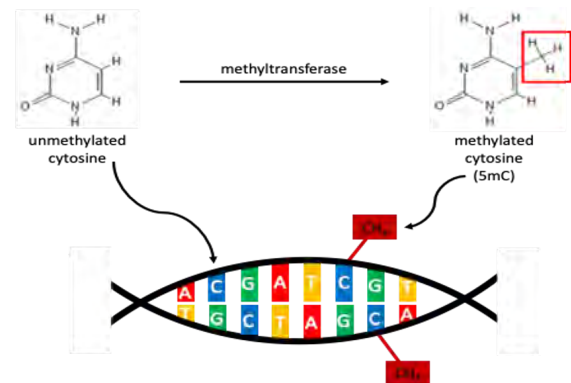


Figure 1. Structure and Mechanism of DNA Methylation. DNA is methylated by methyltransferases which adds a methyl group onto the 5' carbon at CpG sites.

DNA methylation data shows modest correlation with other hallmarks of aging and lifespan: cell senescence, the cessation of cellular division, and telomere attrition, the shortening of the DNA sequence at the ends of chromosomes (López-Otín, et al., 2013). Cellular senescence correlates with changes in DNA methylation at certain loci, which may indicate some degree of interplay between the molecular mechanisms, although the nature of the relationship is not clear (Koch, et al., 2013). However, some hallmarks of aging demonstrate a degree of exclusivity from DNA methylation, such as telomere attrition, as epigenetic age increases even in the presence of telomerase, an enzyme that lengthens telomeres (Kabacik, et al., 2018). The greatest correlation between another hallmark of aging and DNA methylation is seen between stem cell exhaustion and patterns of DNA methylation, as DNA methylation is crucial in maintaining the undifferentiated state of stem cells (Li, et al., 2022).

These correlations between DNA methylation patterns and hallmarks of aging need to be further investigated experimentally to determine the exact degree of molecular mechanism interplay between the processes.

3. Histone Modification

Histone proteins can undergo many types of PTMs, such as phosphorylation, acetylation, and methylation. However, histone acetylation and methylation are by far the most researched (Sterner and Burger, 2000). It is challenging to track histone modification, as it is generally fluid and dynamic in nature and shows less of an overall pattern as it does specific alterations at certain gene loci (Sen, et al., 2016). Research in *D. melanogaster*, *C. elegans*, and *S. cerevisiae* has demonstrated that histone modifications, or histone marks, regulate gene expression through affecting the density of chromatin packaging, either promoting or inhibiting transcription depending on the site of the mark (Booth and Brunet, 2016; Kouzarides, 2007). However, not all mechanisms of histone modification are conserved between model organisms and mammals, increasing the demand for thorough histone modification research in mammalian cell cultures (Dang, et al., 2009). This section of the review covers known histone acetylation and methylation events that are correlated to aging.

3.1 Histone Acetylation

Histone acetylation loosens chromatin folding due to the acetyl group neutralizing the charge of lysine residues in histone side chains, leading to transcriptional derepression (Kouzarides, 2007). Enzymes called histone deacetylases, of which include sirtuins, deacetylate histones (Houtkooper, et al., 2012). A loss of sirtuins has been linked with aging in both *S. cerevisiae* and humans, due to the loss of telomere silencing as histone acetylation increases (Dang, et al., 2009; Michishita, et al., 2008). Global histone deacetylation correlates with increased longevity in *S. cerevisiae* through suppression of oxidative stress responses, as chronic oxidative stress is known to play a role in cell

damage, death, and aging (Eisenberg, et al., 2009). Research in *S. cerevisiae* has also shown that acetylation of lysine 56 on histone H3 (H3K56) at the promoters of histone genes decreases with age, corresponding to decreased gene expression and loss of core histones (Dang, et al., 2009). Conversely, acetylation of lysine 16 on histone H4 (H4K16) in telomeres increases with age, correlating to the loss of silencing in these regions and a shortening of the cell's replicative lifespan, or premature cell senescence (Dang, et al., 2009). Premature cell senescence is associated with the loss of telomere silencing and telomere dysfunction (Michishita, et al., 2008), once again demonstrating interplay of histone acetylation and other hallmarks of aging.

3.2 Histone Methylation

Histone methylation is the addition of methyl groups to arginine, histidine, or lysine residues in histone side chains. Compared to DNA methylation or histone acetylation, histone methylation often results in more versatile effects. Furthermore, multiple histone methylation marks can occur simultaneously on the same histone, which produces a different effect on chromatin remodeling and gene expression than when a mark occurs alone (Greer and Shi, 2012). The effect of methylation may also differ depending on the amino acid methylated (Greer and Shi, 2012).

Lysine methylations, such as trimethylation of histone H3 at lysine 4 (H3K4me3) and lysine 27 (H3K27me3), are modifications that significantly influence transcription and gene expression as aging progresses (Sen, et al., 2016). H3K4me3 is traditionally associated with transcriptional activation, while H3K27me3 is associated with transcriptional repression (Rothbart and Strahl, 2014). Changes in these marks with age can vary widely between cell types and organisms. For example, global H3K27me3 increases in quiescent adult skeletal muscle stem cells yet decreases in somatic *C. elegans* cells with age (Liu, et al., 2013; Maures, et al., 2011). In murine hematopoietic stem cells, H3K27me3 levels mostly increase with age, but some loci show decreased levels of methylation (Sun, et al., 2014). Similarly, studies in *C. elegans* and *D.*

melanogaster show varying patterns of H3K4me3 with age depending on cell type (Booth and Brunet, 2016). Nonetheless, the causal relationship of histone methylation and aging is clear. The ability to methylate and demethylate is a crucial process that decreases with age, where overexpression of enzymes that add methyl groups (methyltransferases) or strip them away (demethylases) has resulted in increased longevity in model organisms, and many of these pathways are highly conserved (Maures, et al., 2011; Greer and Shi, 2012).

It is hypothesized that DNA methylation influences patterns of histone methylation, indicating an interaction between these two types of epigenetic modifications to repress gene expression (Fuks, 2005). However, it is yet unclear which modification occurs first, and further research of this phenomenon may lead to a deeper understanding of the order and progress of aging (Bartke, et al., 2010).

4. Chromatin Landscape Remodeling

Chromatin is wrapped around proteins called histones, forming units of nucleosomes. These nucleosomes can be packaged tightly in a state called heterochromatin, which represses transcription, or packaged loosely in a state called euchromatin, which enables transcription. The genome may display various states of chromatin density depending on loci and the phase of the cell cycle. The heterochromatin loss model of aging presents the loss of tightly packed DNA as a cause of aging due to the increase in genome instability (López-Otín, et al., 2013). Research in *C. elegans* and *D. melanogaster* has demonstrated that maintaining heterochromatin is a conserved process important for muscle strength and longevity (Larson, et al., 2012). Furthermore, heterochromatin was seen to decrease in aging cells and in cells of individuals with Hutchinson-Gilford progeria syndrome (HGPS), a premature aging disease (Greer and Shi, 2012).

Changes in the chromatin landscape associated with aging can be attributed to patterns of DNA methylation and histone modification throughout the genome (Bannister and Kouzarides, 2011). For example, H3K27me3, a mark that induces heterochromatin, is lost on the inactive X

chromosome in HGPS cells (Shumaker, et al., 2006). Additionally, studies have shown that high levels of CpG methylation coincide with heterochromatic regions and repressed gene expression (Fuks, 2005).

Histone methylation may also indirectly contribute to the loss of heterochromatin through the loss of nucleosomes (Booth and Brunet, 2016). The H3K27me3 modification downregulates genes that code for histones and has been linked to the loss of core histones with age (Liu, et al., 2013). This may cause the loss of nucleosome occupancy and the loss of chromatin density, possibly resulting in genomic instability, insertions of transposable elements, and an increase in genomic mutations (Liu, et al., 2013). As mutations accumulate, the susceptibility to age-related diseases increases, leading to decreased health- and lifespan.

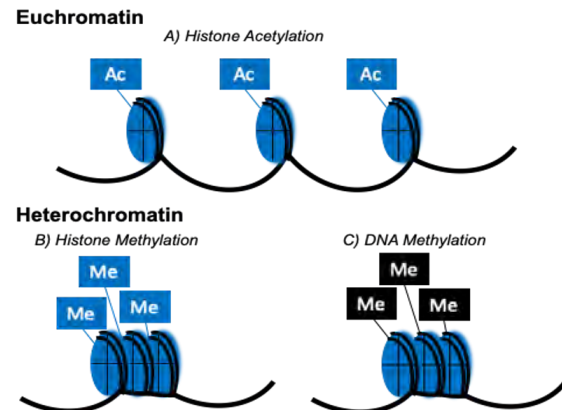


Figure 2. Chromatin Types and the Effects on the Chromatin Landscape. Euchromatin is when DNA is open due to histone acetylation and ready for use via transcription or replication. Heterochromatin is when DNA is tightly wound and inaccessible to proteins important for transcription and replication processes. This DNA tightening is due to histone and DNA methylation.

Similarly, as DNA becomes globally hypomethylated with aging, chromatin becomes less tightly packed (Figure 2), resulting in genomic instability and transposon activation. When heterochromatin unravels into euchromatin, enzymes have easier access to the DNA and gene expression becomes less regulated. Uncontrolled gene expression may lead to the conversion of healthy cells to cancer cells, which is indicative of

accelerated biological age (Horvath, 2013). Epigenetic alterations also result in the loss of heterochromatin at transposable elements, increasing the probability of transposition and consequently mutations that occur due to transposition and result in DNA damage.

Increased histone acetylation also contributes to aging by increasing genomic instability and telomere dysfunction through loss of heterochromatin (Figure 2). Telomeres are ideally maintained permanently in the heterochromatic state outside of DNA replication to maintain genomic integrity, and failure to maintain this state promotes cellular aging (Bannister and Kouzarides, 2011; Michishita, et al., 2008). For example, the loss of SIRT6, a protein that deacetylates lysine 9 histone H3 (H3K9) at telomeres, which results in increased H3K9 acetylation and thus loss of telomeric chromatin density, leads to premature cell senescence (Michishita, et al., 2008). Also, the natural decrease in telomere length with successive cell replications causes DNA damage signaling, which negatively impacts histone expression. The loss of histones, in turn, decreases chromatin density and perpetuates a state of genomic instability (O'Sullivan, et al., 2010). This state of genomic instability can be further perpetuated through DNA methylation and histone modification (Sen, et al., 2016). Thus, epigenetic modifications and other hallmarks of aging may work synergistically to contribute to the molecular causes of aging.

5. Conclusion

Epigenetic modifications demonstrate correlation with various other hallmarks of aging and shape the chromatin landscape of the genome, regulating gene expression and altering phenotypes. While large strides have been made in the field of DNA methylation and its implications on aging, information on histone modification is comparatively lacking. The fluidity of histone modification may offer an easier path toward forced chromatin remodeling and suppression of age-related ailments than may DNA methylation alone, and the coexistence of some histone marks may be a property to take advantage of in the development of targeted

pharmaceuticals. Further research is needed to understand the true relationship between each hallmark of aging and where they show correlation or causation to better comprehend the initiation of cellular decline and to extend healthspan as lifespan increases.

Acknowledgments

The authors would like to acknowledge Lumiere Education for funding this research and the Weill Institute for Cell and Molecular Biology at Cornell University for their resources in the production of this review.

References

- Ashapkin, V. V., Kutueva, L. I., & Vanyushin, B. F. (2017). Aging as an Epigenetic Phenomenon. *Current Genomics*, 18(5), 385–407. <https://doi.org/10.2174%2F1389202918666170412112130>
- Bannister, A. J., & Kouzarides, T. (2011). Regulation of Chromatin by Histone Modifications. *Cell Research*, 21(3), 381–395. <https://doi.org/10.1038/cr.2011.22>
- Bartke, T., et al. (2010). Nucleosome-Interacting Proteins Regulated by DNA and Histone Methylation. *Cell*, 143(3), 470–484. <https://doi.org/10.1016/j.cell.2010.10.012>
- Booth, Lauren N., & Brunet, A. (2016). The Aging Epigenome. *Molecular Cell*, 62(5), 728–744. <https://doi.org/10.1016/j.molcel.2016.05.013>
- Bormann, F., et al. (2016). Reduced DNA methylation patterning and transcriptional connectivity define human skin aging. *Aging Cell*, 15(3), 563–571. <https://doi.org/10.1111/acel.12470>
- Dang, W., et al. (2009). Histone H4 lysine 16 acetylation regulates cellular lifespan. *Nature*, 459(7248), 802–807. <https://doi.org/10.1038/nature08085>
- Day, K., et al. (2013). Differential DNA methylation with age displays both common and dynamic features across human tissues that are influenced by CpG landscape. *Genome Biology*, 14(9), R102.

- <https://doi.org/10.1186/gb-2013-14-9-r102>
- Eisenberg, T., et al. (2009). Induction of autophagy by spermidine promotes longevity. *Nature Cell Biology*, 11(11), 1305–1314. <https://doi.org/10.1038/ncb1975>
- Fuks, F. (2005). DNA methylation and histone modifications: teaming up to silence genes. *Current Opinion in Genetics & Development*, 15(5), 490–495. <https://doi.org/10.1016/j.gde.2005.08.002>
- Garagnani, P., et al. (2012). Methylation of ELOVL2 gene as a new epigenetic marker of age. *Aging Cell*, 11(6), 1132–1134. <https://doi.org/10.1111/accel.12005>
- Greer, E. L., & Shi, Y. (2012). Histone methylation: a dynamic mark in health, disease and inheritance. *Nature Reviews Genetics*, 13(5), 343–357. <https://doi.org/10.1038/nrg3173>
- Hannum, G., et al. (2013). Genome-wide Methylation Profiles Reveal Quantitative Views of Human Aging Rates. *Molecular Cell*, 49(2), 359–367. <https://doi.org/10.1016/j.molcel.2012.10.016>
- Hernandez, D. G., et al. (2011). Distinct DNA methylation changes highly correlated with chronological age in the human brain. *Human Molecular Genetics*, 20(6), 1164–1172. <https://doi.org/10.1093/hmg/ddq561>
- Horvath, S. (2013). DNA methylation age of human tissues and cell types. *Genome Biology*, 14(10), R115. <https://doi.org/10.1186/gb-2013-14-10-r115>
- Houtkooper, R. H., Pirinen, E., & Auwerx, J. (2012). Sirtuins as regulators of metabolism and healthspan. *Nature Reviews Molecular Cell Biology*, 13(4), 225–238. <https://doi.org/10.1038/nrm3293>
- Johansson, Å., Enroth, S., & Gyllenstein, U. (2013). Continuous Aging of the Human DNA Methylome Throughout the Human Lifespan. *PLoS ONE*, 8(6), e67378. <https://doi.org/10.1371/journal.pone.0067378>
- Jylhävä, J., Pedersen, N. L., & Hägg, S. (2017). Biological Age Predictors. *EBioMedicine*, 21, 29–36. <https://doi.org/10.1016/j.ebiom.2017.03.046>
- Kabacik, S., et al. (2018). Epigenetic ageing is distinct from senescence-mediated ageing and is not prevented by telomerase expression. *Aging (Albany NY)*, 10(10), 2800–2815. <https://doi.org/10.18632/aging.101588>
- Kaeberlein, M., Burtner, C. R., & Kennedy, B. K. (2007). Recent Developments in Yeast Aging. *PLoS Genetics*, 3(5), e84. <https://doi.org/10.1371/journal.pgen.0030084>
- Kirkwood, T. B. L. (2005). Understanding the Odd Science of Aging. *Cell*, 120(4), 437–447. <https://doi.org/10.1016/j.cell.2005.01.027>
- Koch, C. M., et al. (2012). Pluripotent stem cells escape from senescence-associated DNA methylation changes. *Genome Research*, 23(2), 248–259. <https://doi.org/10.1101/gr.141945.112>
- Kouzarides, T. (2007). Chromatin Modifications and Their Function. *Cell*, 128(4), 693–705. <https://doi.org/10.1016/j.cell.2007.02.005>
- Larson, K., et al. (2012). Heterochromatin Formation Promotes Longevity and Represses Ribosomal RNA Synthesis. *PLoS Genetics*, 8(1), e1002473. <https://doi.org/10.1371/journal.pgen.1002473>
- Li, A., Koch, Z., & Ideker, T. (2022). Epigenetic aging: Biological age prediction and informing a mechanistic theory of aging. *Journal of Internal Medicine*. <https://doi.org/10.1111/joim.13533>
- Liu, L., et al. (2013). Chromatin Modifications as Determinants of Muscle Stem Cell Quiescence and Chronological Aging. *Cell Reports*, 4(1), 189–204. <https://doi.org/10.1016/j.celrep.2013.05.043>
- López-Otín, C., et al. (2013). The Hallmarks of Aging. *Cell*, 153(6), 1194–1217. <https://doi.org/10.1016/j.cell.2013.05.039>
- Maures, T. J., et al. (2011). The H3K27 demethylase UTX-1 regulates *C. elegans* lifespan in a germline-independent, insulin-dependent manner. *Aging Cell*, 10(6), 980–990. <https://doi.org/10.1111/j.1474-9726.2011.00738.x>
- Michishita, E., et al. (2008). SIRT6 is a histone H3 lysine 9 deacetylase that modulates telomeric chromatin. *Nature*, 452(7186), 492–496. <https://doi.org/10.1038/nature06736>
- O’Sullivan, R. J., et al. (2010). Reduced histone biosynthesis and chromatin changes arising from a

damage signal at telomeres. *Nature Structural & Molecular Biology*, 17(10), 1218–1225.
<https://doi.org/10.1038/nsmb.1897>

Pal, S., & Tyler, J. K. (2016). Epigenetics and aging. *Science Advances*, 2(7), e1600584.
<https://doi.org/10.1126/sciadv.1600584>

Rothbart, S. B., & Strahl, B. D. (2014). Interpreting the language of histone and DNA modifications. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, 1839(8), 627–643.
<https://doi.org/10.1016/j.bbagr.2014.03.001>

Sen, P., et al. (2016). Epigenetic Mechanisms of Longevity and Aging. *Cell*, 166(4), 822–839.
<https://doi.org/10.1016/j.cell.2016.07.050>

Shumaker, D. K., et al. (2006). Mutant nuclear lamin A leads to progressive alterations of epigenetic control in premature aging. *Proceedings of the National Academy of Sciences*, 103(23), 8703–8708.
<https://doi.org/10.1073/pnas.0602569103>

Spólnicka, M., et al. (2017). DNA methylation in ELOVL2 and C1orf132 correctly predicted chronological age of individuals from three disease groups. *International Journal of Legal Medicine*, 132(1), 1–11.

<https://doi.org/10.1007/s00414-017-1636-0>

Sterner, D. E., & Berger, S. L. (2000). Acetylation of Histones and Transcription-Related Factors. *Microbiology and Molecular Biology Reviews*, 64(2), 435–459.

<https://doi.org/10.1128/mmbr.64.2.435-459.2000>

Sun, D., et al. (2014). Epigenomic Profiling of Young and Aged HSCs Reveals Concerted Changes during Aging that Reinforce Self-Renewal. *Cell Stem Cell*, 14(5), 673–688.

<https://doi.org/10.1016/j.stem.2014.03.002>

Tan, Q., et al. (2016). Epigenetic Drift in the Aging Genome: A Ten-Year Follow-up in an Elderly Twin Cohort. *International Journal of Epidemiology*, 45(4), 1146–1158. <https://doi.org/10.1093/ije/dyw132>

Watt, F., & Molloy, P. L. (1988). Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. *Genes & Development*, 2(9), 1136–1143.

<https://doi.org/10.1101/gad.2.9.1136>

Winnefeld, M., & Lyko, F. (2012). The aging epigenome: DNA methylation from the cradle to the grave. *Genome Biology*, 13.

<https://doi.org/10.1186/gb-2012-13-7-165>

Lupus and Multiple Sclerosis: Understanding the Pathophysiology of Under- Funded Autoimmune Diseases and Potential Treatments

Aarushi Mupparti^{1*}, Niki Itharaja¹, Neal Itharaja¹

¹Dublin High School, Dublin, CA USA

Received August 22, 2022; Revised May 1, 2023; Accepted, June 2, 2023

Abstract

The immune system is a body system made up of cells, organs, and proteins that work together to defend the body against foreign invaders. However, there are times when the immune system attacks and kills its healthy tissue and cells, a condition known as autoimmunity. An autoimmune disease occurs when the immune system doesn't know the difference between the body's cells and foreign pathogens, so it attacks its cells. This paper will examine two autoimmune diseases: Lupus and Multiple Sclerosis (MS). Lupus is a chronic autoimmune disease that is often separated into five types. Lupus is mostly characterized by the inflammation it causes throughout the body, including a noticeable facial butterfly rash. Over 5 million people are affected by Lupus and there is no cure for Lupus, but trials have had varying results depending on the severity of the disease, with some treatments including NSAIDs inhibitors. MS is a chronic autoimmune disease characterized by fatigue, impaired vision, and tremors. MS is caused when the immune system destroys the protective sheath of nerve fibers, disrupting the flow of information in the nervous system, and causing nerve damage. There are about 2.3 million people diagnosed with MS. As of 2022, Lupus received over \$139 million in funds while MS received over \$20 million. Although MS and Lupus are both autoimmune diseases that affect a significant population, they do not receive nearly as much funding when compared to diseases like Cancer or Parkinson's. This paper is dedicated to increasing awareness of the widespread effects of autoimmune diseases and the need for more funding. We have compiled background information and statistics using credible search engines and peer-reviewed studies for collective understanding.

Keywords: Autoimmunity, Lupus, Multiple Sclerosis, Funding, Immune System, Treatments, Clinical Trials

1. Introduction

Lupus is one of the hardest diseases to diagnose because of its complexity and how its symptoms overlap with other disorders one common symptom is shown in Figure 1. Therefore, multiple tests must be taken before it is confirmed that a patient has lupus. Put simply, a combination of blood and urine tests, as well as a physical diagnosis, will help identify whether someone has lupus. However, symptoms

vary from person to person, making it hard to know whether someone has lupus.

MS is another disease that is hard to diagnose due to its complexity and varying symptoms. Some of the symptoms include tremors, pain, mood swings, slurred speech, vertigo, and involuntary movement caused by damage to the protective layer of a neuron called Myelin Sheath, as shown in Figure 2. Once the neurologist suspects the patient has MS, they will use a checklist known as the "McDonald criteria" to

* Corresponding Author
aarurose13@gmail.com

Advisor: Gabriel Rodriguez
gabriel.j.rodriguez@vanderbilt.edu

diagnose the patient. Some of the tests include MRIs, Blood tests, Spinal taps, and Evoked potentials. An MRI is used to detect inflammatory lesions in the brain which appear as irregular, white matter on the scan, similar to Figure 3.



Figure 1. Lupus Facial Rash also known as the “Butterfly Rash”

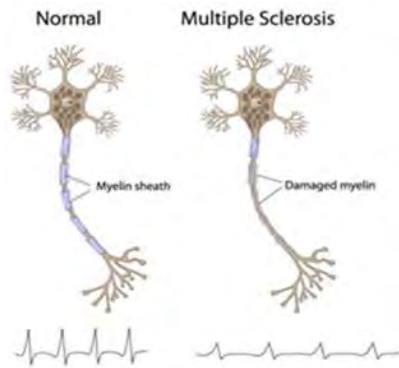


Figure 2. Difference between a normal neuron and a neuron that is affected by Multiple Sclerosis

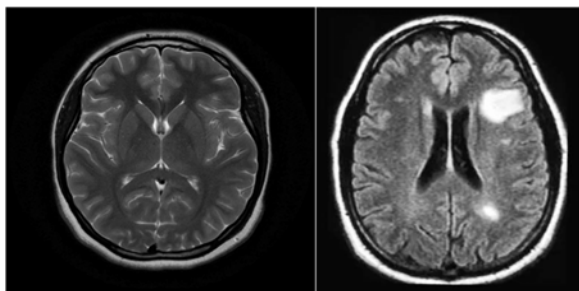


Figure 3. Difference between a normal MRI (left) and a MRI of a Multiple Sclerosis patient (right).

However, the test isn't very reliable because it's not specific, so a patient can receive a false positive and be mistakenly diagnosed with MS. Spinal Taps analyze your spinal fluid to check for elevated spinal fluid protein and white blood cell count. The

elevation in immunoglobulin G (IgG) levels indicates MS as it is a sign of an autoimmune response. The least useful test is the Evoked potentials test as it simply checks for response times to stimuli to indicate if your nerves are damaged, and factors, such as age, can easily affect them.

2. Data Collection Methods

For this research paper, articles from the past 10 years were found using the National Center for Biotechnology Information (NCBI) search engine and EBSCO information center. Words such as autoimmunity, immune system, lupus, multiple sclerosis, diagnosis, tests, etc., were used to find accurate information on the two autoimmune diseases. These keywords were used in the Google search engine to find credible sources, such as the Mayo Clinic. This paper focuses on two rare autoimmune diseases, Lupus, and MS. Lastly, the figures in this paper are taken from articles in which information on the diseases was given or used by NCBI.

3. Results

There has been a significant number of studies and trials done regarding treatments and more efficient management of Lupus. Some clinical trials, studies, and possible treatments have been described. A research study was conducted on the pipeline to automate the classification of glomeruli features in renal biopsy. The researchers concluded that they “propose a complete pipeline for the classification of WSIs into various LN stages based on glomeruli features” (Gupta, et al., 2021). A Pilot Study of CC-220 to Treat SLE (Systemic Lupus Erythematosus) was also done, and results indicated that the drug had improvements to be made but was more successful than the placebo (Celgene, 2020). From an engineering perspective, author Xubin Hao and authors from the Department of Rheumatology and Immunology at the Hospital of Nanjing University Medical School reviewed the engineering technologies and proposed their challenges in SLE treatment. In their review article titled “Developing engineering technologies for the treatment of

systemic lupus erythematosus”, they explain that “Compared with traditional therapeutic methods, nanocarriers have multiple advantages in agents delivery for the treatment of SLE, such as enhanced drug solubility, sustained release, and passive accumulation at inflammatory sites...active targeting of tissues or cells in need of delivery can also be achieved... mainly used for the selective delivery of glucocorticoids and immunosuppressants”. They further discuss other technologies such as CAR-T cell therapy and noted how “the largest sample in the study of CAR-T treatment of lupus is only 5 patients, and further research is needed on larger samples and whether it can treat early or mild SLE”. Lastly, they review the efficiency of nanoparticles which are “encapsulated in RBC membrane and tumor CM has been effectively tested in MRL/lpr mice, but other CM could also be introduced to treat lupus” (Hao et al., 2023). This research study concluded that engineering technologies such as the ones shown in Figure 4 have a high probability of treating and curing SLE in comparison to the relatively static development of drugs.

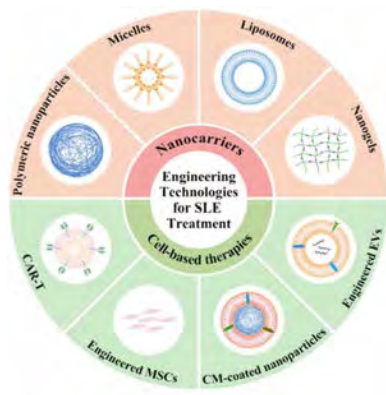


Figure 4: Schematic illustration of recent engineering technologies for Systemic lupus erythematosus (SLE)

There have been many clinical trials and studies that have been conducted for MS as well, and some have promising results. One of them includes a study of Rebif® (interferon beta-1a) in Subjects With Relapsing Multiple Sclerosis. The study with interferon beta-1a was successful and interferon beta-1a can be used to treat MS (Merck KGaA, Darmstadt, Germany, 2018). Another study compared

“Oral Ponesimod Versus Teriflunomide In Relapsing MS” indicated that “Ponesimod” or “PONVORY™” was more successful than Teriflunomide at reducing symptoms in those with relapsing MS (Actelion, 2022). However, Ponesimod was noted to have mild to severe side effects including hypertension and upper respiratory infections. Researcher Masoud H. Manjili from the Department of Microbiology & Immunology at the Virginia Commonwealth University School of Medicine conducted research on a new approach to MS therapy and treatment. In his paper “The adaptation model of immunity: A new insight into aetiology and treatment of multiple sclerosis”, Manjili notes that current treatment for MS consists of immune suppressive interventions because of the self–nonself (SNS) model and the danger model. These models state that “inflammatory immune responses towards the central nervous system (CNS), triggered by breakage of tolerance to self-antigens (SNS model)1 or by damage-associated danger signals triggered by events such as menstrual cycle result in MS” (Manjili, 2023). However, he proposed the adaptation model of immunity which offers a different approach for MS treatment. This theory “suggests that inflammatory immune responses are required for neuronal homeostasis in the CNS as well as restoration of remyelination process in patients with MS... This immunological function is coordinated through the adaptation receptors (AdRs) expressed on the CNS, and their nominal adaptation ligands (AdLs) or co-receptors expressed on immune cells, that is signal IV. I...However, alterations in the expression of the neuronal AdRs could shift neuroprotective T cells into neurodegenerative T cells. This theoretical model can incite new direction for research and drug development for patients with MS” (Manjili, 2023). As the medical field advances over time, new discoveries and experiment trials will yield more promising treatments, and hopefully, a tentative cure.

4. Funding

Compared to other neurological diseases such as ALS or Parkinson’s, MS has the least publication rate in medical journals. A review of publications from 1985 to 1999 showed that MS also had the least

productivity in scientific papers. MS has relatively high funding, as evidenced by The National Institutes of Health (NIH) shown in Figure 6 where MS receives over 20 million dollars annually for research. This is more than typically expected, but the money isn't as equally distributed amongst all aspects of research. The funding is mostly given to the most promising and innovative research proposals. However, another aspect of research in MS that is underfunded is the recruitment of high-achieving and credible researchers who can deliver the most accurate information on this chronic neurodegenerative disease. Lupus gets about 139 million dollars per year in funding as of 2022. While Lupus only gets 139 million dollars in funding, Cancer gets about 6.25 billion dollars of funding from the National Cancer Institute (NCI). \$139 million in funding to dig deeper into Lupus is a good amount compared to the \$90 million given in 2014. Although Cancer is a more common disease and therefore received much more funding, it's evident that Lupus is getting more attention as the medical field is expanding.

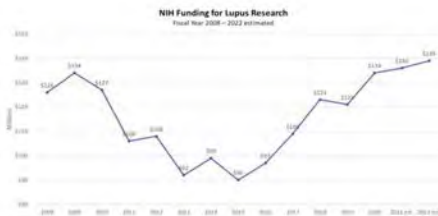


Figure 5: The graph demonstrates the amount of money (in millions) the NIH is funding for Lupus across the United States from 2008-2022.

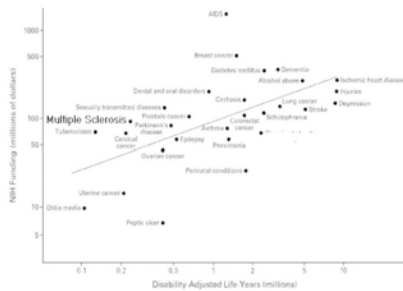


Figure 6: The graph demonstrates the relationship between the amount of NIH funding (1996) and the Disability Adjusted Life Years of patients across 27 different diseases/conditions.

5. Conclusion

Both Lupus and Multiple Sclerosis are difficult to diagnose because there is no single test to assuredly diagnose them. Many of the symptoms of these autoimmune disorders overlap with other conditions and therefore require rounds of testing for a positive diagnosis. Compared to MS, Lupus gets much higher funding.

According to the Lupus Foundation of America, 1.5 million Americans, and around 5 million people worldwide have a form of lupus. It is diagnosed mainly between the ages of 15 - 44 with around 16,000 cases of lupus appearing annually. About 10% - 15% of people with lupus die prematurely due to the disease. Worldwide, about 2.3 million people are diagnosed with MS, while in the US specifically, over 1 million people are affected by it. About 200 cases of MS are diagnosed weekly. As of 2022, lupus receives \$139 million in funding annually as shown in Figure 5, while MS receives only \$20 million for its research. While Lupus has been getting more attention from the medical field, as evidenced by its increase in funding (in 2014, lupus received \$90 million annually), MS has been relatively the same. MS has the least publication rates and productivity rates in scientific journals, especially compared to other neurodegenerative diseases such as ALS or Parkinson's, while its funding is primarily used to look into new, innovative treatments for the disease instead of being used to bring in more credible researchers who can deliver more accurate and useful information on the disease. MS receives less attention than lupus primarily because it is far rarer than lupus. Thus, lupus has a higher public awareness that gives it an edge over MS in receiving funding for its treatments. The social stigma surrounding these two diseases also distinctly separates them. Many see MS as affecting only older people, aged well into their 60s and 70s, while lupus is more prevalent in younger populations. Of course, this isn't true, yet the misconception remains, and so Lupus is regarded as a more pressing issue compared to MS. To solve the disparities between MS and Lupus funding there must be increased public awareness about the two diseases. Eliminating such misconceptions as MS affecting older populations and lupus affecting

younger people, as well as highlighting the need for support for MS patients, among other things, would give the public a more well-rounded understanding of why both diseases need to be addressed seriously and equally.

References

Actelion. (2022, January 26). Oral Ponesimod Versus Teriflunomide In Relapsing Multiple Sclerosis (OPTIMUM). ClinicalTrials.gov. Retrieved June 24, 2022, from <https://clinicaltrials.gov/ct2/show/NCT02425644?term=Multiple+Sclerosis&rslt>

Celgene. (2021, September 28). A Study to Evaluate the Efficacy and Safety of CC-220 in Subjects With Active Systemic Lupus Erythematosus. ClinicalTrials.gov. Retrieved June 24, 2022, from <https://www.clinicaltrials.gov/ct2/show/NCT03161483?term=CC-220>

Chelmow, D. (n.d.). *Diseases & Conditions - Medscape reference*. Diseases & Conditions - Medscape Reference. Retrieved April 28, 2023, from <https://emedicine.medscape.com/>

Department of Defense: Lupus Research Program. Lupus Foundation of America. (n.d.). Retrieved June 23, 2022, from <https://www.lupus.org/lupus-research-program>

Gupta, A., Reddy, A., Jawahar, C., & Vinod, P. (2021, December 14). Classification of histopathology images using ConvNets to detect Lupus Nephritis. *arxiv.org*, 5. Retrieved June 24, 2022, from <https://arxiv.org/pdf/2112.07555.pdf>

Hao, X., Wang, Y., Liu, R., Zhang, D., Kong, B., Liang, J., & Sun, L. (2023). Developing Engineering Technologies for the treatment of systemic lupus erythematosus. *Biomedical Technology*, 4, 1–10. <https://doi.org/10.1016/j.bmt.2023.02.002>

How Lupus Affects the Body. (n.d.). Johns Hopkins Lupus Center. Retrieved June 24, 2022, from <https://www.hopkinslupus.org/lupus-info/lupus-affects-body/>

How is MS diagnosed | Multiple Sclerosis Society UK. (n.d.). MS Society. Retrieved June 24, 2022,

from <https://www.mssociety.org.uk/about-ms/diagnosis/the-tests-for-ms>

JBT-101 in Systemic Lupus Erythematosus (SLE) - Full Text View - ClinicalTrials.gov. (n.d.). ClinicalTrials Retrieved June 24, 2022, from <https://www.clinicaltrials.gov/ct2/show/NCT03093402>

Lupus: Causes, symptoms, and research. (n.d.). Medical News Today. Retrieved June 24, 2022, from <https://www.medicalnewstoday.com/articles/323653#types>

Lupus - Symptoms and causes. (2021, January 27). Mayo Clinic. Retrieved June 24, 2022, from <https://www.mayoclinic.org/diseases-conditions/lupus/symptoms-causes/syc-20365789>

Lupus nephritis - Symptoms and causes. (2022, January 14). Mayo Clinic. Retrieved June 24, 2022, from <https://www.mayoclinic.org/diseases-conditions/lupus-nephritis/symptoms-causes/>

Manjili, M. H. (2023). The adaptation model of immunity: A new insight into aetiology and treatment of multiple sclerosis. *Scandinavian Journal of Immunology*, 97(4). <https://doi.org/10.1111/sji.13255>

Mayo Foundation for Medical Education and Research. (2021, January 27). *Lupus*. Mayo Clinic. Retrieved June 23, 2022, from <https://www.mayoclinic.org/diseases-conditions/lupus/symptoms-causes/syc-20365789>

Mayo Foundation for Medical Education and Research. (2022, January 7). *Multiple sclerosis*. Mayo Clinic. Retrieved June 23, 2022, from <https://www.mayoclinic.org/diseases-conditions/multiple-sclerosis/diagnosis-treatment/drc-20350274>

Merck KGaA, Darmstadt, Germany. (2018, September 20). A Study of Rebif® in Subjects With Relapsing Multiple Sclerosis (RELIEF). ClinicalTrials.gov. Retrieved June 24, 2022, from

<https://clinicaltrials.gov/ct2/show/NCT02064816?term=Multiple+Sclerosis&rslt=With&draw=2&rank=3>

Multiple Sclerosis Association of America – improving lives Today! MSAA. (2023, April 21). Retrieved April 28, 2023, from <https://mymsaa.org/>

NCBI. (2001). *Multiple Sclerosis: Current Status and Strategies for the Future*. National Library of Medicine. Retrieved June 24, 2022, from <https://www.ncbi.nlm.nih.gov/books/NBK222391/>

NCI's role in cancer research. National Cancer Institute. (n.d.). Retrieved June 23, 2022, from <https://www.cancer.gov/research/nci-role>

NCI's role in cancer research. National Cancer Institute. (n.d.). Retrieved June 23, 2022, from <https://www.cancer.gov/research/nci-role>

Radiopaedia. Radiopaedia Blog RSS. (n.d.). Retrieved April 28, 2023, from <https://radiopaedia.org/>

Results. Lupus Research. (2023, April 18). Retrieved April 28, 2023, from <https://www.lupusresearch.org/advocate/advocacy-results/>

Results. Lupus Research. (2022, May 3). Retrieved June 23, 2022, from <https://www.lupusresearch.org/advocate/advocacy-results/>

Rolak, L. A. (2022, September 4). *Multiple Sclerosis: It's Not The Disease You Thought It Was*. NCBI. Retrieved June 24, 2022, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1069023/>

Treating lupus with immunosuppressive medications. Johns Hopkins Lupus Center. (2019, March 27). Retrieved June 23, 2022, from <https://www.hopkinslupus.org/lupus-treatment/lupus-medications/immunosuppressive-medications/>

Scaling an Ensemble ML Algorithm for the Classification of Tree Species Through Satellite Imagery

Neiv Gupta^{1*}

¹Monta Vista High School, Cupertino, CA USA

Received March 12, 2023; Revised April 17, 2023; Accepted, June 9, 2023

Abstract

Dry conditions in the Western United States have increased the frequency and severity of forest fires in the Sierra Nevada Mountain Range. Organizations and cities are actively working towards developing a better understanding of forest structure and dynamics. While tree species classification models in the past have dealt with smaller regions and fewer trees, we hypothesized that it is possible to scale the area and number of trees analyzed by our model without sacrificing model accuracy by adding additional variables to satellite imagery, such as Normalized Difference Vegetation Index (NDVI), Normalized Difference Moisture Index (NDMI), Soil-Adjusted Vegetation Index (SAVI), crown ratio, tree height, and tree diameter. We compared the results of applying the Random Forest (RF) Machine Learning (ML) algorithm to a dataset containing satellite imagery alone and with a dataset containing satellite imagery augmented with object-specific attributes (OSA) such as crown shape, tree height, and tree diameter. We then trained and tested the algorithm across two large and different regions with similar tree species prevalence. After the addition of OSA to training data, the results from the experiment demonstrated a mean classification accuracy increase from 66.4% to 90.2%, thus allowing the ML model to scale over larger regions.

Keywords: Environmental Engineering, Remote Sensing, Forest Management, Machine Learning, Classification

1. Introduction

With the recent increase in frequency, intensity, and duration of forest fire events in the Western United States, cities and researchers are looking to better understand forest structures to prevent and mitigate large fire events. Classifying and mapping tree species provides an efficient and effective way to manage forest inventories and protect forest resources. Accurate maps are also necessary for effectively monitoring drought and fire conditions, which could severely threaten a forest ecosystem (Talukdar, et al., 2020; Ballanti, et al., 2016). These maps could help firefighters better understand a forest's vegetation and characteristics, which are

essential variables to consider when attempting to predict and assess the behavior of an active fire.

Remote sensing is a perfect technique for such tasks, providing synoptic views and information over large areas at very high resolutions. Specifically for tree species classification, remote sensing through high spectral bands of imagery provides the highest resolution and detail for tree species classification. As a result, airborne hyperspectral light detection and ranging (LiDAR) imagery satisfies the optimal conditions for sensors best suited for tree species classification (Immitzer, et al., 2012). LiDAR allows for the capture of spatial patterns of on-the-ground features through multiple spectral bands, which makes it a very useful tool in the field of remote

* Corresponding Author
neiv06@gmail.com

Advisor: Jeff Li Wen
jlwen@stanford.edu

sensing, specifically for the classification of remotely sensed objects. However, airborne LiDAR is not a practical source of imagery due to its high costs and limited availability. As a result, alternative sources of remotely sensed imagery must be considered. Multispectral satellite imagery is a possible alternative to hyperspectral LiDAR imagery, despite its inability to reach the detail and spectral band variety of hyperspectral lidar imagery (Immitzer, et al., 2012; Wang, et al., 2021). The terms hyperspectral and multispectral refer to the electromagnetic spectral band variety of the image. Hyperspectral imagery encompasses more spectral bands, making it more sophisticated than multispectral imagery (Wang, et al., 2021).

The application of machine learning in classification algorithms used in the general field of remote sensing has been increasing in popularity. These algorithms have become increasingly important for general object classification through hyperspectral imagery and multispectral satellite imagery. For example, past research and applications have used the RF machine learning algorithm to classify land cover, map ecological zones and landslides, create forest canopy fuel maps for fire forecasting, and analyze urban tree species inventories (Ballanti, et al. 2016; Immitzer, et al., 2012). In these applications, RF has been used with both hyperspectral data and multispectral satellite imagery because of the large number of input variables provided for the algorithm.

While RF has performed successfully with LiDAR and spectral data in past research (Ballanti, et al., 2016; Ghimire, 2010; Clark and Roberts, 2012), our experimentation demonstrates RF providing 66.4% mean classification accuracy when using satellite imagery alone. We also trained and validated across different regions with similar prevalence of tree species. In this study, we experimented with including OSA, such as crown ratio, tree height, and tree diameter, with satellite imagery to improve the classification accuracy across larger and geologically diverse regions. Our goal was to scale the model by improving model performance over a more extensive area with variations in topographical features and vegetation. We hypothesized that RF would demonstrate higher classification accuracy with the

addition of OSA.

2. Materials and Methods

2.1 Study Area and Data

Our region of study was the Greater Lake Tahoe region/El Dorado National Forest, California (39°58'N, -121°24' W). Our satellite image, downloaded from the United States Geological Survey (USGS) website, was captured from the Landsat 8 Operational Land Imager (OLI). The area is a mix of mountainous terrain and dense temperate forest with elevations ranging from 0 m to 1898 m above sea level, which adds to the significant variance among tree species in the area. Of this larger region, we broke up the dataset into two small subregions, one in the Northern Greater Lake Tahoe Region (R1) and the other in the El Dorado National Forest (R2), south of the Greater Lake Tahoe Region. We optimized our data this way because the scale of the study site was too large to train and cross-validate our machine-learning models. The bounds of the full satellite image also contained non-forested land, such as shrubland, agricultural land, urban areas, etc., which could confound our models and lead to misclassification. In addition, we specifically selected the two subregions in the northern and southern regions of the Greater Lake Tahoe region because it allowed us to see if the models were scalable on very similar, but not identical, regions.

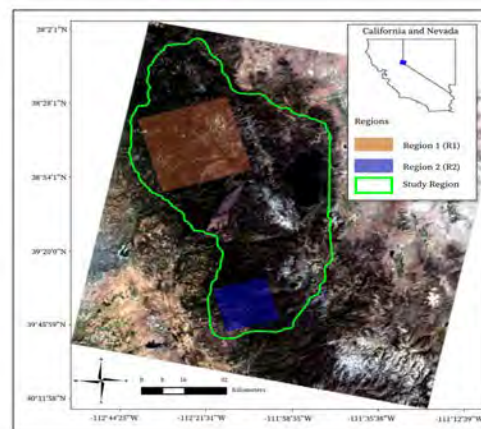


Figure 1. Study site and location of testing regions. The image was acquired from the United States Geological Survey database and captured by the Landsat 8 OLI satellite.

Our ground truth data came from the US Department of Agriculture (USDA) Forest Service TreeMap2016, a tree-level model of forests in the conterminous United States. Both regions have a similar distribution of tree species, with the tree species discussed in our study being the most prevalent in the region. The reason for the distribution of tree species not being identical across both regions is that the distribution of tree species varies due to environmental factors such as altitude, mean temperature, forest density, weather patterns, etc.

Table 1. The distribution of tree species in the regions of study. The table also includes the total number of trees of the selected species in R1 and R2. The data were imputed from the Forest Inventory and Analysis database, which the USDA Forest Service matched onto a raster grid. We processed the raster data of our study regions and computed the distributions for the most prevalent tree species in the regions.

Scientific Name	R1 Count		R2 Count	
	('000)	%	('000)	%
<i>Abies concolor</i> (AC)	377.1	24.3	84.1	23.9
<i>Arbutus menziesii</i> (AM)	30.5	2.0	17.7	22.7
<i>Calocedrus decurrens</i> (CD)	179.3	11.6	148.6	13.6
<i>Cornus nuttallii</i> (CN)	68.7	4.4	11.1	13.5
<i>Pinus lambertiana</i> (PL)	148.5	9.6	63.3	10.2
<i>Pinus ponderosa</i> (PP)	88.5	5.7	24.5	4.4
<i>Pseudotsuga menziesii</i> (PM)	467.6	30.2	140.9	4.0
<i>Quercus chrysolepis</i> (QC)	125.3	8.1	84.7	3.0
<i>Quercus douglasii</i> (QD)	25.4	1.6	27.5	2.9
<i>Quercus kelloggii</i> (QK)	38.3	2.5	18.7	1.8
Total	1,549.4	100.0	621.0	100.0

2.2 Pre-processing and Data Formatting

The satellite imagery we used for our analysis were the second band (blue, 0.450 - 0.51 μm), third band, (green, 0.53 - 0.59 μm), fourth band (red, 0.64 - 0.67 μm), fifth band (near infra-red (NIR), 0.85-0.88 μm) and sixth band (Short-wave infrared (SWIR1), 1.57-1.65 μm). NDVI, NDMI, and SAVI were computed using NIR and SWIR1. Crown Ratio, Tree Height, and Tree Diameter were obtained using the

USDA TreeMap2016 data set.

We also re-projected the satellite imagery on the WGS84 coordinate reference system (CRS) to match the TreeMap2016 raster image's CRS, NAD83 Conus Albers. Due to the bounds of the full satellite image encompassing non-forested land and our system running into image processing constraints, we cropped the dataset to the two further subregions within the larger image.

We balanced our datasets using random undersampling to prevent data imbalances and an uneven dataset. Random undersampling balances an uneven dataset by keeping all data points in a minority class and decreasing the size of the majority class to match the size of the minority class. The data points removed from the majority classes are chosen randomly (Hasanin and Khoshgoftaar, 2018).

2.3 Classification

For model training, we used a cross-validation approach. We trained our models on Region 1 and validated on Region 2 (train-test pair of R1, R2), and trained on Region 2 and validated on Region 1. We performed this experiment using satellite imagery alone as well as satellite imagery with OSA data. The inputs for our models were NDMI, NDVI, SAVI, the strength values of the red, green, and blue bands of satellite imagery represented as a 16-bit digital notation, as well as OSA data using the USGS TreeMap2016 dataset.

For our classification, we applied the RF machine learning algorithm. RF is a non-parametric ensemble learning algorithm consisting of a large number of decision trees, which enhances traditional decision trees. An individual bootstrapping sample (sampling with replacement) is utilized to construct each decision tree. At each node of the tree, the split determination is based on the Gini criterion. With standard decision trees, nodes are split by the variable that provides the best split or the highest decrease in Gini. However, RF randomly selects a subset of variables at each node and chooses the best splitting variable. New data are classified from a majority vote among the classification outcomes of all constructed decision trees. For determining a rough estimate of the classification error, the out-of-bag data (OOB),

the samples not in the bootstrapping sample, are used. Each decision tree is used to classify the samples with the OOB dataset. Finally, for each sample in the original data set, the majority vote of the corresponding decision trees is compared with the truth labels, resulting in an estimate of the misclassification rate (Immitzer, et al., 2012; Breiman, 2001). For our model, we set our parameters such that warm_start=False, n_estimators=100, and max_depth=100. These parameters ensure the algorithm uses adequate decision trees with significant depth to increase the robustness of the model and improve model performance.

3. Results

For each train-test pair for both satellite imagery and satellite imagery + OSA, we calculated the classification accuracy, precision, recall, F1-score, and Cohen’s Kappa coefficient. We then constructed a confusion matrix for each train-test pair for satellite imagery + OSA to determine the tree species with the highest mean classification accuracies among all algorithms. Model precision indicates the accuracy of the model in terms of how many instances that the model classified as a certain label were actually correct. Model recall indicates the accuracy of the model in terms of how many instances were correctly classified over the total number of instances for that specific label. The F1-score is the harmonic mean of precision and recall. Kappa reflects the model’s true accuracy without the addition of correct classifications due to random chance (Yacouby and Axman, 2020).

For training and validation on satellite imagery + OSA, our model exhibited a mean classification accuracy of 90.20%, mean precision of 90.90%, mean recall of 90.14%, mean F1-Score of 90.14% and mean Kappa of 89.07%.

Table 2. Percentage Accuracy Metrics Table, Satellite Imagery + OSA

Train-Test	Accuracy	Precision	Recall	F1-Score	Kappa
(R1, R2)	95.17	95.75	95.07	95.09	94.53
(R2, R1)	85.23	86.04	85.2	85.18	83.6

For training and validation on satellite imagery alone, our model exhibited a mean classification accuracy of 66.40%, mean precision of 69.04%, mean recall of 66.25%, mean F1-Score of 65.72% and mean Kappa of 62.50%.

Table 3 (all figures in %): Accuracy Metrics Table, Satellite Imagery Only

Train-Test	Accuracy	Precision	Recall	F1-Score	Kappa
(R1, R2)	73.52	75.5	72.46	71.72	69.38
(R2, R1)	59.28	62.58	60.04	59.72	55.61

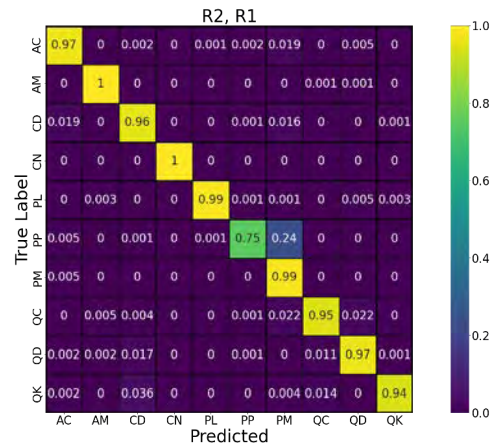


Figure 2. Confusion Matrix: Classification using satellite imagery + OSA. Training-Testing R1/R2

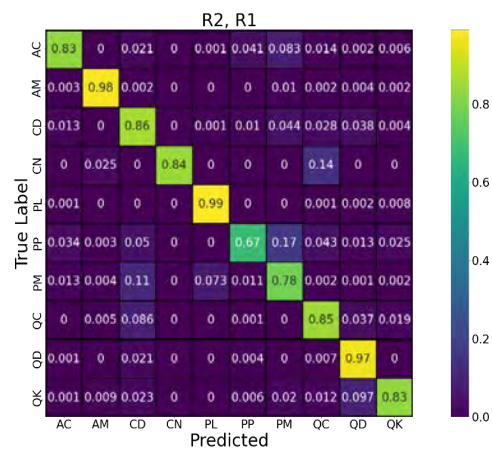


Figure 3. Confusion Matrix: Classification using satellite imagery + OSA. Training-Testing R2/R1

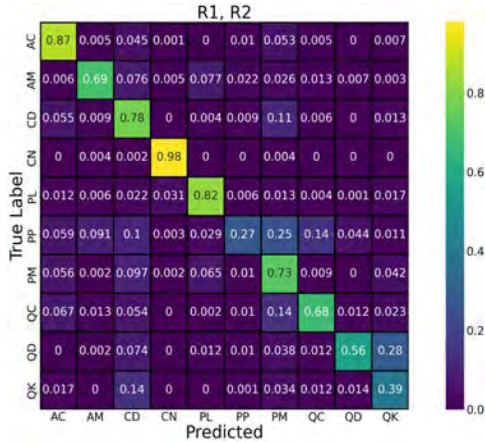


Figure 4. Confusion Matrix: Classification using satellite imagery only. Training-Testing R1/R2

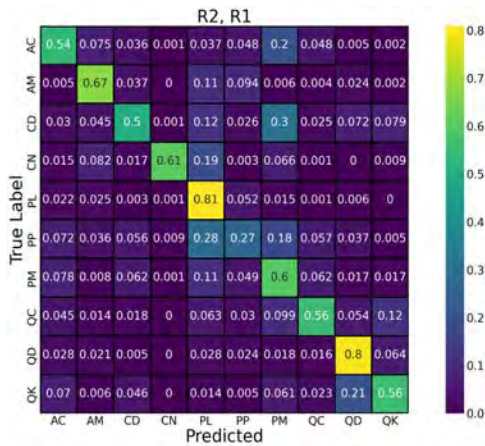


Figure 5. Confusion Matrix: Classification using satellite imagery only. Training-Testing R2/R1

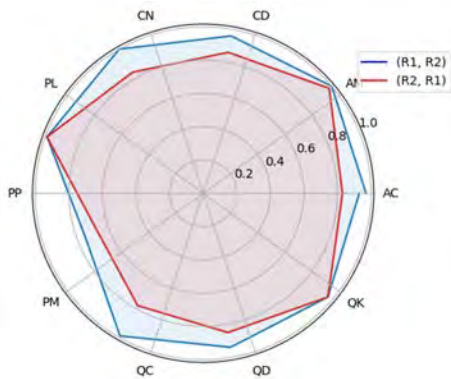


Figure 6. Spider chart comparison of satellite imagery + OSA tree species-specific classification accuracy for both Train-Test Pairs. The spider chart compares RF’s overall performance for each Train-Test pair. The distance of an algorithm’s polygon’s edge to the end of the spoke reflects the accuracy the algorithm demonstrated for that specific Train-Test pair.

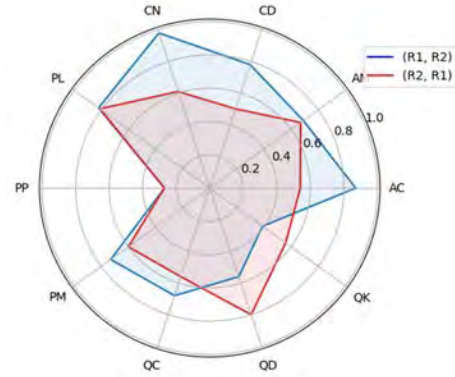


Figure 7. Spider chart comparison of satellite imagery only tree species-specific classification accuracy for both Train-Test Pairs.

4. Discussion

The addition of object-specific attributes to satellite imagery significantly improved classification accuracy across regions. Tables 2 & 3 show a summary of the performance metrics of the RF classification for cross-region Training and Testing split (Train-R1:Test-R2, Train-R2:Test-R1), under both scenarios, using satellite imagery alone and satellite imagery augmented with OSA. When adding OSA, the average classification accuracy increased from 66.4% to 90.2%. In addition, the Kappa values went from indicating moderate to strong agreement between the classification results and reference data by increasing from 62.50% to 89.07%. When comparing the confusion matrices in Figures 2 & 3 to the confusion matrices in Figures 4 & 5, the percentage of correct classifications in Figures 2 & 3 is higher across all tree species than those in Figures 4 & 5. The comparison between Figures 6 & 7 reflects the same observation, as the distance between the labels and edges of the R1, R2 & R2, R1 polygons in Figure 6’s spider chart is significantly less than that of Figures 7’s spider chart, indicating higher tree species-specific classification accuracy for satellite imagery + OSA. Overall, introducing OSA in our models improved scalability by increasing model accuracy when training and validating on separate regions. The difference in accuracy metrics between train-test pairs is possibly due to variations in the geographical features or environmental dynamics between regions. In addition, the disparity in total instances, displayed in

Table 1, could potentially explain the difference in accuracy metrics, as Region 1 having more instances than Region 2 allows the model to possess more knowledge when training on Region 1 and validating on Region 2 than it does when training on Region 2 and validating on Region 1, thus yielding a higher classification accuracy for the R1, R2 train-test pair.

There could be many underlying reasons for reduced classification accuracy when using satellite imagery values only. This includes a minimal distinction between RGB strength values pixels between different tree species, a complex forest structure of the study region, and a top-viewed pixel-based classification approach for tree species based on a large plot of land. However, adding OSA provides better distinction for RF algorithms to iterate, resulting in better classification accuracy.

When processing the data for analysis, we ran into numerous memory issues because of the size and scale of our datasets. We initially tried to encompass a significantly larger cutout of the Greater Lake Tahoe/El Dorado National Forest region to ensure ample geographic diversity but continuously ran into errors because of our system's limitations on memory. To bypass this issue, we experimented with incremental learning and k-fold cross-validation as possible solutions, but our system continued to run into memory issues. Especially for studies concerning large datasets, incremental learning allows the model to be trained from a series of batches, compared to the entire dataset at once, which could pose issues depending on the strength of the system used for data analysis and processing. Specifically, incremental learning is learning through streaming data, which arrives over time without sacrificing the model's accuracy. As a result, the models' overall accuracy when training and validating different general regions could potentially have improved with a stronger system designed for handling larger datasets and a successful implementation of incremental learning. A stronger system could potentially process a larger study area with higher OSA specificity, allowing models to encompass larger regions with more geographical diversity without sacrificing performance. Further research that implements incremental learning, k-fold cross-validations, and a stronger processing system

could potentially help construct more sophisticated models that are more accurate and encompass larger regions.

5. Conclusion

Mapping tree species provide an effective way to manage forest inventories and resources. While high classification accuracy for tree species is possible for small regions using satellite imagery, this research concludes that scaling of RF ML algorithm across a wider region is possible with high classification accuracy by including OSA such as crown shape, tree diameter, and tree height to satellite imagery. Since this research focused on the Greater Lake Tahoe region, additional investigations should explore the applicability of these findings in other regions and introduce incremental or k-fold cross-validation approaches to further improve model performance and scalability.

Acknowledgments

I would like to thank my mentor, Mr. Jeff Wen (Stanford School of Earth, Energy & Environmental Sciences), for his guidance throughout this project. I would also like to thank Dr. Jeff Liu (MIT Lincoln Laboratory) for helping me build foundational knowledge for this research and Mr. Ryan Malhoski (GIS Manager, City of South Lake Tahoe) for helping identify potential datasets.

References

- Ballanti, L., et al. (2016). Tree Species Classification Using Hyperspectral Imagery: A Comparison of Two Classifiers. *Remote Sensing* 8(6) 445
<https://doi.org/10.3390/rs8060445>
- Breiman, L. (2001). *Machine Learning* 45(1), 5–32.
<https://doi.org/10.1023/a:1010933404324>
- Clark, M. L., & Roberts, D. A. (2012). Species-Level Differences in Hyperspectral Metrics among Tropical Rainforest Trees as Determined by a Tree-Based Classifier. *Remote Sensing* 4(6), 1820–1855.
<https://doi.org/10.3390/rs4061820>
- Ghimire, B., Rogan, J., & Miller, J. (2010). Contextual land-cover classification: incorporating

spatial dependence in land-cover classification models using random forests and the Getis statistic. *Remote Sensing Letters* 1(1), 45–54. <https://doi.org/10.1080/01431160903252327>

Hasanin, T., & Khoshgoftaar, T. (2018). The Effects of Random Undersampling with Simulated Class Imbalance for Big Data. *IEEE International Conference on Information Reuse and Integration (IRI)*. <https://doi.org/10.1109/iri.2018.00018>

Immitzer, M., Atzberger, C., & Koukal, T. (2012). Tree Species Classification with Random Forest Using Very High Spatial Resolution 8-Band WorldView-2 Satellite Data. *Remote Sensing* 4(9), 2661–2693. <https://doi.org/10.3390/rs4092661>
Talukdar, S., et al. (2020). Land-Use Land-Cover

Classification by Machine Learning Classifiers for Satellite Observations—A Review. *Remote Sensing*, 12(7), 1135. <https://doi.org/10.3390/rs12071135>

Wang, Y., et al. (2021). Classification of Street Tree Species Using UAV Tilt Photogrammetry. *Remote Sensing* 13(2), 216. <https://doi.org/10.3390/rs13020216>

Yacoub, R., & Axman, D. (2020). Probabilistic Extension of Precision, Recall, and F1 Score for More Thorough Evaluation of Classification Models. *Proceedings of the First Workshop on Evaluation and Comparison of NLP Systems, Association for Computational Linguistics*, 79-91. [doi:10.18653/v1/2020.eval4nlp-1](https://doi.org/10.18653/v1/2020.eval4nlp-1).