



Silent Inheritance: Evaluating the Transgenerational Effects of Epigenetic Modifications

Bhargava Kanneganti

Introduction

In the past 5 years, the field of epigenetics, a field of science that studies how environmental factors can influence DNA expression, has changed to show the effects of environmental factors in inducing disease spread and death in future generations through parent offspring inheritance.¹ The main environmental factors that induce these epigenetic changes are pollution, diet, and chronic stress which all have shown individually through current research to lead to epigenetic changes, and the respective transgenerational spread of diseases and death rates in future generations.¹ Although most current research studies the impact of these environmental factors individually, in the real world these factors hardly ever single handedly induce epigenetic changes in individuals. Especially in regions like the Southeastern United States, industrial pollution, lack of proper access to nutritious foods and diet education, and historical racial and resulting socioeconomic problems altogether affect individuals in these regions.² Because of this, it is important to research how all of these environmental factors together influence transgenerational epigenetic changes compared to single influences to understand the health of the current and future populations in the southeast region of the US.³ This research study aims to find the combined effects of pollution, diet, and chronic stress on causing epigenetic changes, and their resulting influence of disease spread in current and future populations of the Southeast region of the US.

To find the strength of these relationships, this research conducted correlational analysis on current datasets and studies through secondary sources, from the past decade. Then data was organized into categories for each quantitative variable (pollution, diet, stress, and resulting epigenetic changes), and then analyzed it through linear regression. Analysis of all of the environmental factors together showed that having more environmental factors present showed a greater combined correlation on inducing epigenetic changes that resulted in greater disease spread that could pass onto future generations.

These correlations show that researching and targeting the combined effects of multiple environmental factors can help reduce disease spread in future generations, while also helping researchers find cures for epigenetic induced diseases in the current generation.¹ These correlations also emphasize the importance of future multi-environmental factor epigenetic research, and the need for public health support both from policymakers and healthcare professionals in the Southeast region of the US, not just to help reduce disease spread in the current population, but to also prevent disease inheritance in future populations.

Literature Review

Epigenetics is a field of science that studies how gene expression is affected by different environmental factors without changing the original DNA of an organism. In epigenetics, gene expression changes because of factors such as DNA methylation, the addition of methyl groups to cytosine bases,⁴ and histone modifications, changes to the histone proteins that determine how tightly or loosely DNA is packed. Based on traditional Mendelian inheritance, DNA sequences have always been the only factor that determines an offspring's inheritance of genes.⁵ However, continued research into epigenetics shows that environmental factors can cause changes in gene expression across multiple generations. Early research of epigenetics focused on the belief that epigenetic changes could not be inherited, but recent research by Tim Burton and Neil Metcalfe, evolution biologists studying how early life conditions influence future generations,³ and by Edith Heard and Robert Martienssen, geneticists who specialize in

transgenerational inheritance mechanisms, has proved that epigenetic changes actually do carry on across multiple generations in both animal and human populations.⁶

These transgenerational influences especially appear in regions like the Southeastern United States, where industrial pollution, certain diets, and continuing socioeconomic challenges paired with chronic stress combine to create the greatest risk of epigenetic changes, and their associated increased chances of disease and death. Robert Feil and Mario Fraga, researchers who have examined how environmental patterns influence epigenetic variation, emphasize the importance of understanding these factors together.⁷ Johannes Bohacek and Isabelle Mansuy, neuroscientists studying the nongenetic inheritance of new behaviors, argue that finding these epigenetic marks can bring up intervention sessions that can improve long term health outcomes.⁸

Epigenetics & pollution

For a long time, researchers have linked environmental pollution to harmful health effects, but only in recent decades have they begun to link these effects to epigenetic changes. Mustafa Syzf, an epigeneticist who researches how environmental exposures influence gene regulation and behavior,⁹ and Feil and Fraga, who study rising patterns in epigenetics and the environment, suggest that DNA methylation is an important process in which environmental pollutants can change gene activity.⁷ Xiaoji Wu and Yi Zhang, molecular biologists researching DNA demethylation and its mechanisms, have proved how methylation marks can stay long after the first exposure from the environment, leading to increased susceptibility to chronic diseases.¹⁰

Current research suggests that organisms facing harsh environmental conditions may go through epigenetic changes to survive. In human populations, exposure to particulate matter (PM2.5) or heavy metals can cause methylation changes that pass across generations. Andrea Baccarelli and Valentina Bollati, public health researchers examining epigenetics and environmental chemicals, emphasize that these chemicals can impact respiratory, metabolic, and immune function epigenetically.¹¹ However, opinions still vary on how permanent these changes actually are. Rachel Yehuda et al, who study the continued epigenetics presence in Holocaust survivors, emphasize the cases where high environmental stress has caused epigenetic marks to continue.¹² On the contrary, Szyf, Heard, and Martienssen suggest that improving environments might slowly reverse some epigenetic changes, emphasizing the possible plasticity, or adaptability of the epigenome. A study in which urban populations were exposed to high PM2.5 levels showed a lower amount of methylation markers after pollution levels decreased, showing that the epigenome could be reversed.¹¹ Furthermore, while some experts suggest that air pollution is the main cause of transgenerational epigenetic changes, others such as Baccarelli and Bollati indicate that waterborne contaminants like arsenic/cadmium are also as important.¹¹ These arguments show the complicatedness of the problem, with different types of pollutant, lengths of exposure, and the different genes of every person all working together to cause epigenetic changes.

Epigenetics & diet

While dietary effects on health have long been a problem, only recently has research found that epigenetic mechanisms can now show how bad nutrition can create new epigenetic marks on genes. Kaspar Daniela and other nutrition researchers studying transgenerational epigenetic inheritance,¹³ and Baccarelli and Bollati, who study environmental epigenetics, all

suggest that a diet of high sugars and fats can lead to DNA methylation changes.¹¹ These changes can affect metabolic and inflammatory pathways, the pathway in which diseases develop, and can last not just in the person making those diet choices, but also in their offspring. Although, still, researchers do not always agree on the duration of these diets caused epigenetic marks.

While some research suggests that constantly eating high fat/sugar diets can create significant changes to methylation patterns throughout one's entire life,¹³ Burton and Metcalfe's studies, which emphasize the impact of early life conditions,³ and Feil and Fraga's research, which suggests that environmental effects can actually be changed,⁷ prove that improving access to healthy foods and specific nutrition treatments could slowly change these epigenetic patterns. These ideas suggest that diet caused epigenetic changes are neither permanent nor completely reversible. However, they remain a continuing problem that public health officials might slowly help reduce through better policies, health education, and diet support/access.

Epigenetics & chronic stress

Chronic stress caused by socioeconomic difficulties, discrimination, or past trauma, not only affects the mental health of a person, but it can also create new epigenetic marks. Yehuda et al, known for their research on Holocaust survivors and their offspring, have shown how trauma can change methylation patterns in the genes that regulate stress, while possibly causing chances of anxiety, depression, and reduced immune system function.¹² On the other hand, Szyf's research suggests how these stress related epigenetic changes can occur in entire communities that are facing constant difficulties, emphasizing the long term passing down of these epigenetic changes across generations.⁹

Still, not all scholars agree on the everlastingness of these stress caused epigenetic changes. But, Heard and Martienssen suggest that epigenetic modifications are rapidly changing compared to previous years,⁶ while Burton and Metcalfe suggest that improving the environment can decrease the epigenetic effects on future generations.³ Instead of seeing these epigenetic stress markers as irreversible, some researchers suggest that through entire community support, better socioeconomic conditions, and proper therapy, communities can help change epigenetic patterns back to normal, especially for future offspring.

Research Gap

All of these findings show how environmental factors such as pollution, diet, and chronic stress can cause harmful changes in gene expression throughout one's entire life, and in his/her offspring. In the Southeastern United States, industrial pollution, poor diets, and social difficulties that lead to continued stress, all must come together for researchers to fully understand the reason behind the epigenetic changes in this region. Feil and Fraga's suggestion of cleaning up and fixing environmental pollution and causes of mental stress,⁷ could help slowly reverse the possibility of epigenetic changes in future generations, and their associated higher chances of disease and death. Although researchers have studied how pollution, diet, and stress individually affect the epigenome, most research to this date has only looked at these factors separately. Only a few studies have looked at how all of these factors work together, especially in the Southeast region of the US, where these factors are commonly found together when causing epigenetic changes. This gap in the literature creates the need to research the combined epigenetic influence of multiple environmental factors, which could possibly show important epigenetic effects that researchers cannot identify when researching them individually.

Accordingly, the research question arises: How do pollution, diet, and chronic stress together influence transgenerational epigenetic changes and disease susceptibility in the Southeastern region of the US?

Methods

This study used correlational research to determine how pollution, diet, and stress impact epigenetic changes across multiple generations. Correlational research was chosen for this study because the research question meant to find the relationship between environmental factors and epigenetic changes. Since the goal of this study was to find if pollution, stress, and diet correlated with increased chances of developing epigenetic changes, a correlational methodology was chosen to find statistical patterns between existing samples of data, without having to do direct experiments. In addition, this research design was chosen as experimentation with real life human populations and the long term monitoring of constantly changing environmental factors to find their impact on epigenetic changes would not be ethically possible, nor feasible in the time frame. Instead, by analyzing already existing secondary source data, correlational research could help find significant relationships between the number of environmental factors and their resulting epigenetic changes, ethically and efficiently.

Many studies influenced the methodology of this research study, with each source having a correlational research design in common, but different methods of sourcing, collecting, and organizing data. Breton et al's study researched the correlation between exposure to air pollution and the resulting DNA methylation patterns in children, finding that increased exposure to air pollutants led to increased chances of gaining epigenetic changes and their resulting higher chances of respiratory diseases.¹⁴ However, their study used longitudinal data collected over many years through experimentation and focused primarily on the effects of pollution, while this research focuses on the combined effect of multiple environmental variables (pollution, stress, and diet) on epigenetic changes and uses more recent and short term data. On the other hand, Heijmans et al studied how maternal malnutrition during the Dutch Hunger Winter correlated with transgenerational epigenetic changes being passed onto their offspring.¹⁵ Their study used historical data collected many years after the famine, while this research design primarily uses more recent data to find the effects of modern day environmental factors on epigenetic changes. In addition, Heijmans et al research focused only on the effects of diet, compared to the multiple environmental factors that this study focuses on.¹⁵

Collecting the data

The first step was gathering secondary sources that had data on pollution, stress, diet, and their effects on epigenetic changes. Studies were found through reputable databases such as Google Scholar and PubMed using specific search terms such as "DNA methylation AND pollution," "epigenetics AND diet," "transgenerational effects AND stress," and vice versa to make sure that only and as many as possible relevant studies were used. Reputable databases were determined based on their credibility in the field of epigenetics research, their number of citations, and the amount of peer-reviewed studies that they had. This was done as peer-reviewed studies were usually the most credible, due to their numerous reliable authors, and thus meant that their respective databases were also reputable. Only peer-reviewed studies that were published within the past 10 years were selected to make sure that the findings showed current changes and data in epigenetics, as previous research before this time period focused mainly on the history and development of epigenetics rather than current implications

on modern individual and societal health. Furthermore, studies focusing on human models were primarily used in comparison to mammalian model based studies, as they gave the most useful and reliable data for understanding how environmental factors affect human epigenetic inheritance and could be used to model how these environmental factors and epigenetic changes would play in real life human scenarios and generational inheritance.

After looking through multiple research studies, five sources were selected based on their relevance to the research question and the quality and amount of data that they had.³ Most of the numerical data was directly taken and extracted from tables, figures, and graphs within these peer-reviewed papers and health agency reports, including those published by the Centers for Disease Control and Prevention (CDC), the National Institutes of Health (NIH), and state-level public health databases.¹⁶⁻²⁰ Each dataset was then double-checked and inputted into a master spreadsheet that was cleaned, labeled by variable, and organized into consistent rows across tabs for pollution, stress, diet, and epigenetic changes (refer to appendix A, B, C, and D to see secondary data). Each source was then rechecked for credibility, to again make sure that only factual and reliable sources were used. This step was necessary to make sure that the secondary sources and their data were credible and transparent, since a common limitation of correlational research is that secondary sources' data may at times be fabricated or outdated, causing unintended biases in one's research.

Organizing & cleaning the data

Once the sources were gathered, their data was then sorted and organized into a spreadsheet, where each variable, pollution, stress, diet, and epigenetic changes was assigned its own tab and columns for later data analysis. Sorting of data was important as it would allow for the efficient comparison of environmental factor data with its respective epigenetic change data while preventing syntax and organizational bugs in data from occurring due to disorganized data management.

Pollution data was categorized based on the age of the individual, pollution type (airborne chemicals, heavy metals, etc), source (industrial, vehicular, etc), level of exposure, and associated health outcomes (disease and death rates). On the other hand, stress data included the age of the individual, level of stress, cause (psychosocial, environmental, etc), and associated health outcomes (disease and death rates). While, diet data focused on diet type (healthy, poor, etc), causes of diet (education, resources, etc), and associated health outcomes (disease and death rates). Epigenetic changes were also recorded in a separate tab based on the number of environmental factors present, associated disease and death risks, and the specific environmental factor(s) that caused the changes. These column descriptors were used mainly as means to show data trends of age groups affected, types, and causes of certain environmental factors for later implications, with the diseases and death rates of each environmental factor along with the resulting effects on epigenetic changes being the primary sources of data for data analysis.

Before moving onto analysis, a final data check and debugging were done to identify and fix any numerical or syntax errors, with each point of data being double checked with its source to make sure it was accurate and prevent any numerical errors. During this check redundant data was removed, to prevent shifting the data analysis to a skewed, incorrect correlation. One limitation of this method found during this process was that not all studies contained complete data for each variable leading to gaps in the spreadsheet for both the causes and disease and death rate columns of each environmental factor. Since cause data was intended only for

implications and not data analysis, cause data was left alone, while rows of data that lacked complete death and disease rate data were rechecked with their sources, and were removed from the spreadsheet and analysis if no additional data was found. This was important as even small gaps in the spreadsheet could lead to flawed epigenetic and environmental factor correlations in data analysis.

Data analysis

After data collection, statistical analysis was performed using linear regression on the remaining 20 rows of data per environmental factor. This form of regression was chosen because it is the most common in correlation research, especially in finding the strength and direction of relationships between multiple variables, making it perfect for finding patterns between multiple environmental factors and their combined effects on epigenetic changes. Through linear regression, the Pearson correlation coefficient (r-value) was calculated for each factor. Pollution, stress, and diet were each separately analyzed in comparison with their respective disease and death rates to find each factor's correlations. From there disease and death rates were averaged between the 20 rows of data per environmental factor and were then compared to their respective epigenetic factor count through the calculation of another Pearson's correlation coefficient to find correlations between epigenetic factor count and mean disease and death rates of factors.

Results

After performing Pearson's correlation analysis, the data showed both positive and negative correlations. Contrary to my original hypothesis, pollution had a weak negative correlation with disease rates -0.178 and nearly zero, yet negative correlation with death rates -0.055. These correlations likely suggest that although pollution does have an insignificant impact on causing death rates, in terms of disease spread, factors such as healthcare access and pollution prevention are already available to individuals in the Southeast (SE) US, due to a long history of industrialization in this area, causing associated disease rates to be negative and inverse. Because of this long term history, changes in policy may have resulted in a weakened relationship and presence of pollution to epigenetic changes in the SE US, causing there to be less induced epigenetic diseases due to pollution, and thus lower resulting death rates. I found that stress, and diet all had positive correlations with disease and death rates, but with different levels of strength. Stress displayed the highest correlation with disease rates at +0.290, showing a strong positive relationship between increased exposure to racial and socioeconomic challenges and resulting chances of disease. Diet followed with a positive but weaker correlation value +0.149, suggesting that while these environmental factors can lead to disease risk, the relationships aren't very strong. Death rates, on the other hand, had weaker and negative correlations with environmental factors. Pollution had the lowest negative correlation at -0.055, while diet and stress had the highest negative correlations at -0.262 and -0.284, respectively. These lower, and negative correlations between pollution, stress, diet and their death rates are likely due to modern advancements in medical treatments that help decrease chances of death, even with high chances (or nonsignificant) of disease that are correlated with these environmental factors.

When analyzing the combined effects of all three environmental factors on epigenetic changes and their resulting effects on disease and death rates, I found a weak positive correlation of 0.003 between the number of epigenetic factors (pollution, stress, and diet) and

higher disease rates. In addition, these environmental factors' correlation with death rates remained negative at -0.286. This again suggests that while multiple environmental factors together increase the susceptibility of diseases in individuals (although to a limited degree), modern medical innovations have helped reduce their consequences of death in both current and future populations. While showing how environmental factors do have a positive, although weak correlation with inducing epigenetic changes and their resulting disease rates. Figure 1 on the next page represents all of these correlations in a visual representation, with r-values closer to 1 representing more positive correlations, and r-values closer to -1 representing more negative correlations.

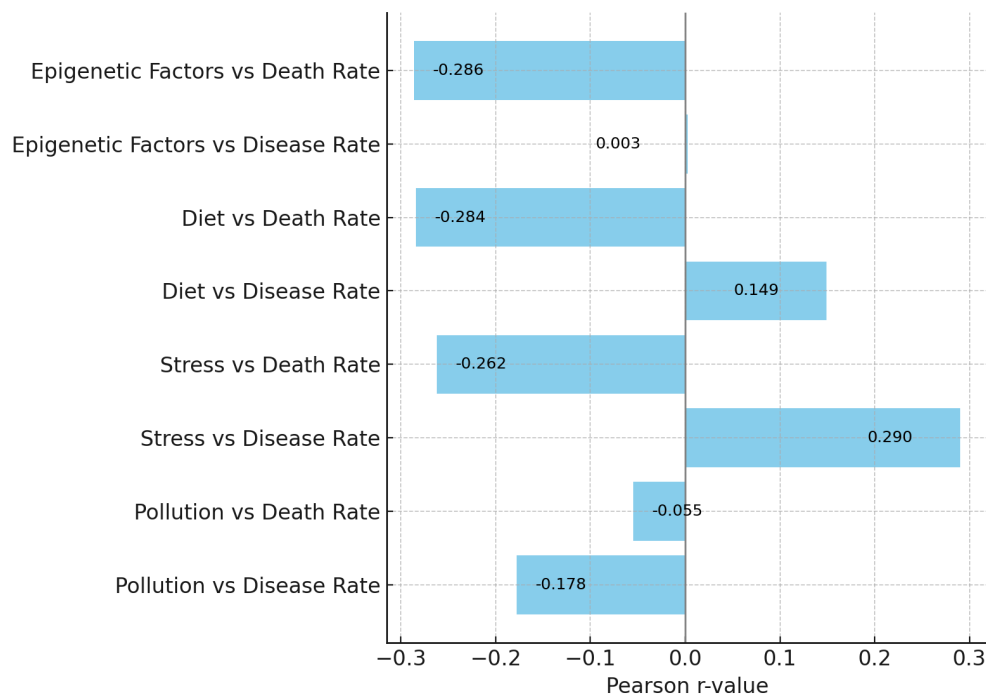


Figure 1 Above. Correlations between pollution, stress, diet, epigenetics, in terms of disease spread and death rates.

Discussion & Analysis

The main goal of this research study was to answer the question: How do pollution, diet, and chronic stress together influence transgenerational epigenetic changes and disease susceptibility in the Southeastern region of the US? My studies results showed that while all three environmental factors, pollution, stress, and diet, cause epigenetic changes and health outcomes, the strength of their correlations varied. Pollution showed a weak negative correlation with disease and nearly no correlation with death rates, while stress and diet showed weak positive correlations with disease and negative or near-zero correlations with death. This may be due modern medical innovations which have allowed to combat and improve the long term outcomes of many diseases.

Understanding the data

This studies data analysis started by conducting a Pearson's correlation coefficient test to find the correlation and strength between environmental factors, epigenetic changes, and resulting disease and death rates. I chose correlation analysis through Pearson's correlation coefficient, so I could find the relationship between multiple quantitative variables (pollution, stress, diet, and epigenetic changes), while not having to do experiments and alter certain variables, which would not be possible due to ethical limitations of performing experiments on human populations and time constraints. After conducting Pearson's correlation coefficient for all of my variables, pollution was found to have the lowest and negative correlation with disease rates, while stress and diet had important but weaker positive correlations. When these environmental factors were analyzed together in comparison to resulting epigenetic changes, these factors showed a combined correlation of 0.003 with epigenetic changes and disease rates. While their combined correlation with death rates and epigenetic changes was low and negative at -0.286, again supporting the idea that although these environmental factors have positive chances of leading to disease (to a limited degree as the correlation is closer to 0), modern medical innovations have minimized the chances of death of individuals with these diseases. Also, by looking at the correlations between each environmental factor and resulting epigenetic changes, disease and death rates, stress was again found to have the strongest and positive correlation, showing how important minimizing stress in the Southeast region of the US is in preventing possible diseases and deaths in current and future populations.

I could have done additional statistical tests, to support my findings even further, such as doing different statistical tests under the category of linear regression to see if there were any additional variables that may have had a presence in slightly skewing my research and its resulting correlations. However, due to time constraints and the lack of lots of data, I chose only Pearson's correlation coefficient because even by itself it gave strong and accurate results and correlation between all four quantitative variables.

Background & limitations

Although everything was done to reduce the limitations of my research and its results, there are still some limitations that were unavoidable due to data and time constraints. One big limitation came from this research study's use of secondary source data, which always has the chance to have biases from the original author, which may end up shifting my results in a negative way. Although, I did all I could by repeatedly rechecking sources for credibility, either by checking if they came from reputable databases, or if they were peer reviewed, it is still not possible for my research to be completely accurate, as even at times inaccurate data from secondary sources could pass through the checks from peer reviews or databases. In addition, another limitation I found was that because this study took data from many sources, there were a lot of problems between how each quantitative variable was measured. For this to not affect my results and its generalizability, I made sure to convert each of the measurements of each variable into a common unit, so that I could easily compare all of my data and create uniform results. Additionally, another limitation I found was that not all of my data is able to represent the entire population of the US, although it can apply to the Southeast region of the US, there are still limits to how much of the population in this region can be represented.

My studies' results of the combined correlation between pollution, stress, and diet to positively influence epigenetic changes in individuals and their resulting disease and death rates, has also been found in other research studies from my discipline. For example, Breton et

al's³ study found that increased air pollution in regions, and increased exposure led to higher chances of DNA methylation in immune system genes, which led to higher chances of developing Asthma in children. My study supports Breton's study, by again developing evidence to the limited but present correlation between pollution exposure and epigenetic changes. Similarly, Heijmans et al's¹⁵ researched the effects of poor diets in pregnant women during the Dutch Hunger Winter, which they found to lead to increased chances of epigenetic changes and linked metabolic and cardiovascular diseases in their offspring. The results of my study again match this finding, showing that poor nutrition in mothers has a high correlation with increased epigenetic changes and resulting disease rates both in the current and future generations. However, my study researched the combined correlation between multiple environmental factors towards influencing epigenetic changes, while both of the previous studies researched the single effect of one environmental factor in influencing epigenetic changes, examined multiple environmental factors simultaneously, showing how more environmental factors have a stronger presence, but although still weak, in causing greater chances of epigenetic changes.

Implications for field of epigenetics

In addition, the results of my study are important especially for general public health, as long as the field of epigenetics. Through my research, finding the correlation between certain environmental factors and epigenetic changes will allow both everyday individuals, epigenetic researchers, and eventually even medical workers to find new solutions and the factors that actually cause certain everyday diseases. From my research and through additional research into the influence of multiple environmental factors in increasing chances of epigenetic changes and diseases, communities across the world can develop new methods to help decrease the prevalence of these environmental factors (specifically pollution, stress, and diet) in society, especially to help decrease chances of the diseases caused by these environmental factors and epigenetics.⁵ Furthermore, my results show the importance of governments, communities, and even families and their need to create their own policies and strategies to decrease their collective chances of diseases due to epigenetic changes, by doing things such as improving nutrition and diet, reducing stress through breaks and fun activities, and using masks and creating mandates to reduce pollution from vehicles and industries to prevent the future damage caused by exposure to pollution.

Future research

Finally, future research should be focused to fix the limitations and gaps of my research study. One important sector of my research study that needs to be researched further in the future, is the long term effect of multiple environmental factors across multiple generations. While my study used other secondary sources that were primarily based on short term experiments, using long term studies and even conducting long term experiments in the future will help find how environmental factors and their resulting increase in epigenetic factors can be passed on over future generations, even when DNA is rechecked for mutations. Lastly, future research done based on my study, should try to take data and conduct experiments on different regions of the US, or even different regions of the world, with larger data samples, to increase the applicability of my research and the correlation between an increased amount of environmental factors, and their combined effects on influencing epigenetic changes and their resulted disease and death rates.

Conclusion

This research study found how pollution, diet, and chronic stress come together to cause epigenetic changes and increase the spread of diseases both in the current population and transgenerationally in the Southeast United States. Through correlation analysis of secondary source data for each of the quantitative variables, this study found that stress had the strongest correlation with inducing epigenetic changes, while diet had weaker, but still positive relationships in inducing epigenetic changes and their resulting disease spread. On the other hand, pollution had a negative but almost nonzero relationship with disease spread, showing the positive effect of pollution reducing policymaking in the SE US. When all of these environmental factors were analyzed together, the combined effect of multiple environmental factors led to a positive but still weak correlation with causing epigenetic changes and higher disease rates, showing the significance of researching the combined effects of environmental factors especially in modelling real world scenarios. These weak, but still positive correlations are beneficial especially in terms of future generation health, because if current populations have a greater chance of getting disease now, then they also have a higher chance of passing on these diseases to their offspring. This will cause a never ending cycle of death and disease spread in all upcoming generations, that will only affect society negatively.

These findings show the significance and provide a future direction for more multi-environmental factor influence based research in the field of epigenetics. With more research focusing on the combined effects of multiple factors, instead of just focusing on the past dominant study of a single factor inducing epigenetic changes. Although there has already been done research on the effects of multiple environmental factors on epigenetics, no study has focused on the specific effects of pollution, stress, and diet especially in the Southeast (SE) region of the US. This research and its findings may not apply to everybody in the SE US, as only certain age groups above the age of 20 were considered, with some age groups and certain subregions of SE US being more present than others in my data. But these findings can still be generalized to individuals above the age of 20 in the SE US to a limited degree, but more importantly these findings show the need for public health support and policy changes to prevent further disease spread in the SE US population. Support may include community health support through diet education workshops, policies to reduce industrial pollution, while also providing access to healthy food options, and mental health support.

Finally, the findings of this research show the importance of researching environmental factors together, to not only help medical professionals develop cures for and to limit the spread of diseases induced by resulting epigenetic changes in the current population, but also prevent these individuals from passing on these diseases to their offspring. The significant impact that epigenetics and the environment can have in influencing the health of current and future generations, should make it clear the necessity of investing in our environment to prevent future harm caused by epigenetic changes.