IJHSR International Journal of High School Research

December 2022 | Volume 4 | Issue 6
ijhighschoolresearch.org
ISSN (Print) 2642-1046
ISSN (Online) 2642-1054

GENIUS OLYMPIAD

"Let's build a better future together"

www.geniusolympiad.org

International Environment Project Fair For Grades 9-12

Rochester, New York

Hosted by Rochester Institute of Technology

RIT Rochester Institute of Technology

@GeniusOlympiad

terra science and education

Table of Contents

December 2022| Volume 4| Issue 6

Editorial Board

International Journal of High School Research

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Comparison of Machine Learning Methods for Early Diagnosis of Parkinson's Disease from Voice Recordings

Aadya Bhat
Eastlake High School, 400 228th Ave N.E., Sammamish, WA, 98074, USA; aadyabhat2005@gmail.com

ABSTRACT: Parkinson Disease (PD) is a common neurodegenerative disease that affects dopaminergic neurons, resulting in muscle impairment. However, in about 25% of patients, PD is incorrectly diagnosed, leading researchers to seek new methods of diagnoses. Specifically, a common side effect of PD is vocal dysfunction, which can be observed up to 5 years before clinical diagnosis. Thus, researchers more recently have focused on speech pattern analysis using machine learning (ML) for improved accuracy and early diagnosis. Using an open-source dataset from University of California Irvine's ML repository, a comparative analysis was conducted on 4 ML techniques -- the Support Vector Machine (Radial Basis Function and Linear Kernel), Multi-Layer Perceptron, and Random Forest. Out of 700+ features available, the tunable q-factor wavelet transform (TQWT) and mel-frequency cepstral coefficient (MFCC) features were extracted. Various combinations of these features were used with the different ML methods to identify the optimal feature and method for increased accuracy in PD diagnosis. Results identified the Random Forest methodology used with half TQWT and MFCC features as the overall best method, with a mean accuracy rate of 86% and area under ROC curve (AUC) measure of 0.86. This is higher than the accuracy rate of clinical diagnosis made by both experts in movement disorders (79.6%) and nonexperts (73.8%).

KEYWORDS: Computational Biology and Bioinformatics; Computational Neuroscience; Parkinson's Disease; Support Vector Machine; Multi-layer Perceptron; Random Forest.

Introduction

Parkinson's Disease (PD), a nervous system related movement disorder that worsens over time,[1] is the second most common neurodegenerative disease in the United States.[2] Additionally, it is projected that the number of PD cases will rise as the global population ages.[3] PD is caused by the death of dopamine secreting neurons (i.e., dopaminergic neurons); by the time of clinical diagnosis, up to 50% of dopaminergic neurons could be lost.[4] The speed at which this loss occurs means that it is too late to significantly slow down the progression of the disease, and treatment options are limited and often ineffective. For this reason, early diagnosis through new diagnostic markers is critical. One novel method for early diagnosis is from analysis of speech patterns.

Along with tremors and other symptoms, speech disorders can manifest in PD cases and have been observed as early as five years prior to clinical diagnosis.[5] Furthermore, in the earlier stages of the disease progression, 90% of patients exhibit vocal dysfunction.[6] However, in about 25% of patients, PD is diagnosed incorrectly,[7] and there is only a 73.8% accuracy with clinical diagnosis made by a non-specialist in movement disorders.[8] Thus, researchers consider machine learning techniques important tools, which can use vocal recordings to learn and implement pattern recognition for improved accuracy in diagnosis. This can be used to solve either regression (finding optimal function) or classification tasks (categorizing data), for early detection of PD by learning to separate audio samples into PD cases or Healthy Control (HC) categories. Within the application of PD diagnostics, this would be a binary classification task, meaning the grouping of data into one of two categories (PD or HC).

Literature Review

In prior research, results of studies using machine learning to detect PD have been undermined due to use of small datasets. Typically, the datasets used have less than 60 PD cases with various successes. For example, research conducted by Tsanas *et al.* (2012) resulted in 98.6% detection accuracy in a dataset with 33 PD patients.[9] Another experiment by Hemmerling *et al.* (2016) using 50 PD and 50 HC had 82% accuracy in males and 90% for females.[10] While these scores are impressive, the small size of the data and lack of clarity in whether validation schemes (such as the leave-one-out cross validation) were applied leaves much room for doubt about generalization to new patients.

While Vaiciukynas *et al.* (2017) address the identified issues of small sample size, the study is conducted with patients speaking in the Lithuanian language and so may not generalize to English speaking patients. Eighteen audio feature sets were extracted from both acoustic cardioid and smartphone microphone audio recordings, then fed through a Random Forest (RF) algorithm involving decision trees with classification tasks. Out of bag error rate, which is a method for measuring predicted error used in accuracy evaluation and tuning of parameters,[2,3] functioned as the validation scheme for classification accuracy. One of the findings of Vaiciukynas *et al.* (2017), was that mel-frequency cepstral coefficient (MFCC) in unvoiced modality (unvoiced segments of signals) was the most important feature for smartphone recordings.[11]

DOI: 10.36838/v4i6.1

Another methodology utilized to solve this classification task was Support Vector Machine (SVM), as used by Hariharan *et al*. (2014). This methodology fed raw dysphonia features (audio features) for feature weighting (assigning higher weights based on importance of feature during selection process), which then was processed for feature selection/reduction, and then finally was classified by least-square support vector machine (LS-SVM), probabilistic neural network (PNN) and general regression neural network (GRNN).[12] This process resulted in 100% accuracy in classification; however, this research used a small dataset from around twenty PD patients, and did not specify validation techniques used. Additionally, overfitting is likely to be an issue when using many different techniques together for a small sample size of twenty PD patients. Therefore, it makes it difficult to state whether this technique is more effective than the previously discussed RF method.

Furthermore, the paper that created the open-source dataset that is used in this paper utilized a different method, Multi-Layer Perceptron (MLP). This research by C.O. Sakar *et al*. (2017) involved over 700 features including tunable q-factor wavelet transform features (TQWT) and MFCC.[13] Overall, prior research has focused on applying or improving a specific machine learning technique for early diagnosis of Parkinson's Disease in patients. Research by C.O. Sakar *et al*. (2017) compared the various individual methods to determine which is more effective and yields a higher classification accuracy. It also compared the effects of 84 MFCC and 432 TQWT on classification accuracy in a top-down approach, removing said features and running analysis on the remaining 238 other features.[13] As noted by Solana-Lavalle *et al*. (2020), various combinations of a smaller selected feature range of between eight to twenty feature sets showed increased accuracy (between 92.68% - 98.4%) compared to research by C.O Sakar *et al*. (2017).[14]

Accordingly, this study's research would add to the existing knowledge by discussing which of the three methods is better for this application in a bottom-up approach, conducting analysis on just the TQWT and MCCF features. This could help identify a minimal set of optimal features and machine learning methods to use for early detection of PD and other diseases that affect vocal production.

Methods

The python coding language was used while working with the various machine learning methods. Python uses a combination of high-level data structures that is both a simple and readable object-oriented approach to programming. Python is considered an ideal language for scripting and application development across many areas in most platforms.[20]

The dataset used is an open-source dataset published by University of California Irvine as a machine learning repository. To produce the dataset, audio samples of sustained 'a' vowel phonation from 252 subjects (188 PD and 64 healthy controls, or HC) were run with various speech signal processing algorithms to extract 754 different features/attributes.[15] Three recordings were taken (in the same sitting) from each subject, totaling 756 recordings or instances in the dataset. In this study, each recording was instead treated as an individual data entry.

Some of the clinically important information extracted for use include data from the application of MFCC and TQWT. MFCCs are described as emulating the filtering properties of the human ear to detect subtle changes in context such as speech identification and recognition, as well as medical diagnostics.[9,13,16] TQWTs are wavelet transforms that are characterized by the tunable q-factor. Wavelet transforms, at a basic level, can be used to decompose data, functions, or operators into different frequency components which are then studied at a scale that matches the resolution. A q-factor is the ratio of center frequency to the bandwidth of the wavelet. The tunable q-factor means that an optimal time frequency representation can be obtained.[13,17,18] Table 1 describes these attributes in context for this application.

Table 1: List and description of the relevant features utilized, as defined and created by C.O. Sakar *et al*. (2017).[13]

Feature	Abbreviation	Explanation	Number of Features
Tunable Q-Factor Wavelet Transform	TQWT	Useful for signal manipulations due to the properties (Q-factor) that allows for decomposition and reconstruction.	432
Mel Frequency Cepstral Coefficients	MFCCs	Used to identify the effect of PD on vocal tract.	84

One of the methods for completion of this binary classification task is the Support Vector Machine (SVM). The SVM method is particularly known for solving linear and nonlinear classification problems, function estimation, and pattern recognition. The SVM essentially utilizes certain kernel functions to map the selected features on the optimal hyperplane, or high dimensional feature space. Here, the SVM available was used as a pre-existing built-in function in scikit-learn (called sklearn, a free software machine learning library). The SVC (support vector classification) is imported from sklearn. SVC input has two arrays, one is x (which is holding the samples of MFCC or TQWT features), and y (of class labels - PD or HC coded as 1 or 0, respectively). This model was trained and validated using the train/test datasets and can then be used for predictions. For different linear and nonlinear (RBF) SVM classification, the kernel can be changed and set to the appropriate type of function, and the method will still remain the same.[19,20,22]

The Random Forest (RF) methodology can also be used for binary classification. RF utilizes decision trees to complete the task. A singular decision tree can be visualized similar to the branches or roots of trees and plants. Decision trees essentially split at nodes, where based on observations and the criteria for each branch of the split, one of the pathways will be chosen. This cycle will repeat with later sub-branches until only homogeneous nodes are left. Random forest is made of a forest of these decision trees that operate as an ensemble, with each of these trees spitting out a class prediction and the class with most votes becoming the model's prediction. In this application, the built-in random forest function from sklearn is used with all the defaults to produce results.[21,22]

Another machine learning technique that can be used is the Multi-layer Perceptron (MLP). MLP is a non-linear mapping of input and outputs, and is made up of input, output, and 1 or more hidden layers with neurons stacked together. MLP is also a type of feed forward neural network, where each layer feeds information forward to the next layer for computation. For the learning aspect, MLP uses backpropagation, which allows for the adjusting of weights through iterations with the goal of loss function reduction (a measure of error). The sklearn library again has a built-in function, which is used with all the default hyperparameters, to classify subjects as PD or HC.[22]

In order to determine the best methodology and feature combination, the following feature sets have been tested for each type of machine learning technique listed above: all TQWT features, half of the TQWT features, all MFCC features, half of the MFCC features, all of the MFCC and TQWT features, and half of the TQWT plus half of the MFCC features. The half sampling in each case was determined by alternating through the available features, selecting every other feature to be used.

Results and Discussion

Results:

The cross-validation technique used was the Stratified K Fold (SKF), which is available in the sklearn library. SKF was used because it is useful for dealing with imbalanced data and smaller datasets. K Fold cross validation involves splitting the data into a specified number of folds (10 in this case), then using the first K folds for training and testing on the rest of the fold. This process is repeated for each fold, and then the mean of the performance metric for all folds is returned.

Results were measured in the form of the Area Under the Curve (AUC) of the Receiver Operating Characteristic Curve (ROC curve). ROC curve is a plot created by graphing the true positive rate against the false positive rate at different threshold settings. AUC is the computed area under the ROC curve and is used for performance evaluation of the machine learning method - the closer to 1.0, the better the performance.

Based on comparisons of the mean AUC metric after cross validation, the best method across all feature sets was RF. As seen in Figure 1, across all methods, the mean AUC for RF is always higher than 0.8. Furthermore, Figure 2 highlights that for all methods (except RF), the highest AUC was achieved with the MFCC feature combination. The feature and machine learning method combination with the overall highest mean AUC, at 0.86, was RF used in conjunction with the half TQWT and MFCC features combination.

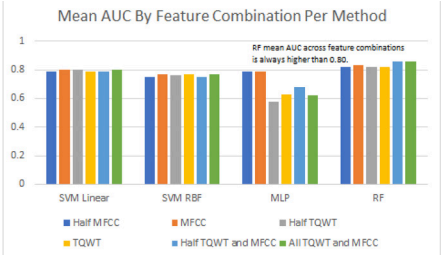


Figure 1: Graph comparison of the mean AUC amongst the different types of feature combinations across machine learning techniques.

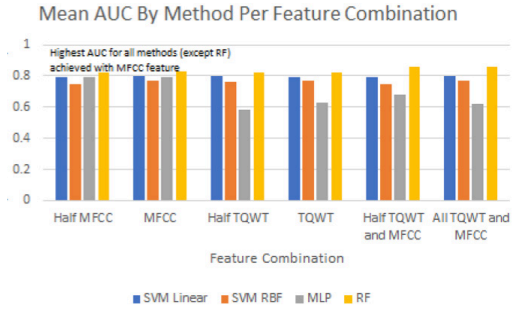


Figure 2: Graph comparison of the mean AUC, set up by method per feature combination.

Discussion:

Across all methods, the classification performance got better when more data (features) was added, except in the case of MLP with the half TQWT and MFCC and all TQWT and MFCC feature combinations. This could perhaps indicate that the data is too high dimensional for the built-in function defaults and could require a more complex network for computation.

The ROC curve plots shed some light on the classification split of different methodologies. As seen in Figure 3, MLP at times is struggling to complete the classification task, with a mean AUC close to that of chance (0.5) when used with half TQWT features. This could be due to dimensionality issues and may be solvable by identifying optimal hyperparameters. MLP also displays a lack of consistency, with a mean AUC range of 0.58 to 0.79 across all feature combinations.

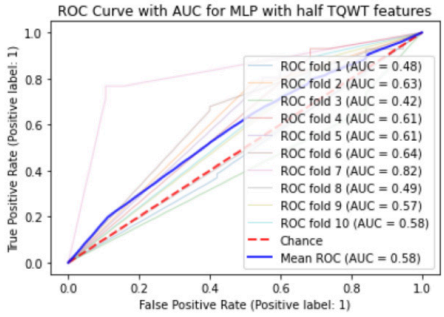


Figure 3: ROC Curve for MLP using the half TQWT features. The mean AUC is 0.58, which is close to the AUC for chance (0.5).

The ROC curve plots for SVM Linear across all feature combinations shows the most consistency, with the mean AUC range being between 0.79 and 0.8. Additionally, the classification split seen in Figure 4 shows an interesting phenomenon: that the SVM Linear and MLP seem to find a similar split. This could potentially signal a legitimate split, with this third class being of "pre-Parkinson's Disease" subjects exhibiting symptoms similar to those seen in PD patients, providing future directions for further development of current labels to include a middle ground.

DOI: 10.36838/v4i6.1

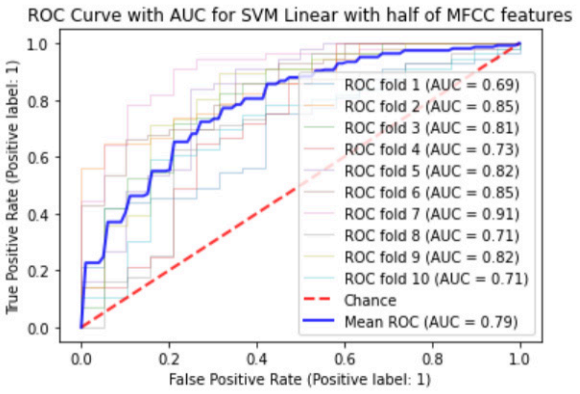


Figure 4: ROC Curve for SVM Linear using the half MFCC features. This split was similar to the MLP split for half MFCC features, and both had the same mean AUC of 0.79.

Conclusions

In conclusion, RF methodology and Half TQWT and MFCC feature set in conjunction resulted in the best AUC measure. Results can be further extended to address observations made including, changing the built-in functions from the defaults or using a more complex network, accounting for a third class of 'Pre-Parkinson's Disease' patients, and addressing potential dimensionality issues. Regardless, this research supports prior findings that the use of voice recordings and machine learning techniques results in higher diagnostic accuracy compared to accuracy of clinical diagnoses. This study aims to contribute to finding a minimal set of vocal features and simple machine learning techniques to detect PD earlier than currently possible and ultimately, through early diagnosis, improve quality of life for PD patients.

Acknowledgements

I would like to acknowledge and thank my mentor, Brianna Marsh, who's support and guidance made this project possible.

References

1. Fadil, R.; Huether, A.; Brunnemer, R.; Blaber, A. P.; Lou, J.-S.; Tavakolian, K. Early Detection of Parkinson's Disease Using Cen ter of Pressure Data and Machine Learning. A*nnu Int Conf IEEE Eng Med Biol Soc* **2021**, *2021*, 2433–2436. https://doi.org/10.1109/EMBC46164.2021.9630451.
2. Beitz, J. M. Parkinson's Disease: A Review. *Front Biosci (Schol Ed)* 2014, 6, 65–74. https://doi.org/10.2741/s415..
3. de Rijk, M. C.; Launer, L. J.; Berger, K.; Breteler, M. M.; Dartigues, J. F.; Baldereschi, M.; Fratiglioni, L.; Lobo, A.; Martinez-Lage, J.; Trenkwalder, C.; Hofman, A. Prevalence of Parkinson's Disease in Europe: A Collaborative Study of Population-Based Cohorts. Neurologic Diseases in the Elderly Research Group. *Neurology* **2000**, *54* (11 Suppl 5), S21-23.
4. Kordower, J. H.; Olanow, C. W.; Dodiya, H. B.; Chu, Y.; Beach, T. G.; Adler, C. H.; Halliday, G. M.; Bartus, R. T. Disease Duration and the Integrity of the Nigrostriatal System in Parkinson's Disease. *Brain* **2013**, *136* (Pt 8), 2419–2431. https://doi.org/10.1093/brain/awt192.
5. Harel, B.; Cannizzaro, M.; Snyder, P. J. Variability in Fundamental Frequency during Speech in Prodromal and Incipient Parkinson's Disease: A Longitudinal Case Study. *Brain Cogn* **2004**, *56* (1), 24–29. https://doi.org/10.1016/j.bandc.2004.05.002.
6. Tjaden, K. Speech and Swallowing in Parkinson's Disease. *Top Geri atr Rehabil* **2008**, *24* (2), 115–126. https://doi.org/10.1097/01.TGR.0000318899.87690.44.
7. Tolosa, E.; Wenning, G.; Poewe, W. The Diagnosis of Parkinson's Disease. *The Lancet Neurology* **2006**, *5* (1), 75–86. https://doi.org/10.1016/S1474-4422(05)70285-4.
8. Rizzo, G.; Copetti, M.; Arcuti, S.; Martino, D.; Fontana, A.; Logro scino, G. Accuracy of Clinical Diagnosis of Parkinson Disease: A Systematic Review and Meta-Analysis. *Neurology* **2016**, *86* (6), 566–576. https://doi.org/10.1212/WNL.0000000000002350.
9. Tsanas, A.; Little, M. A.; McSharry, P. E.; Spielman, J.; Ramig, L. O. Novel Speech Signal Processing Algorithms for High-Ac curacy Classification of Parkinson's Disease. *IEEE Trans Biomed Eng* **2012**, *59* (5), 1264–1271. https://doi.org/10.1109/TBME.2012.2183367.
10. Hemmerling, D.; Orozco-Arroyave, J. R.; Skalski, A.; Gajda, J.; Nöth, E. Automatic Detection of Parkinson's Disease Based on Modulated Vowels. In: *Proceedings of the 17th Annual Confer ence of the International Speech Communication Association (IN TERSPEECH)*. San Francisco, USA; 2016. p. 1190–1194..
11. Vaiciukynas, E.; Verikas, A.; Gelzinis, A.; Bacauskiene, M. De tecting Parkinson's Disease from Sustained Phonation and Speech Signals. *PLoS One* **2017**, *12* (10), e0185613. https://doi.org/10.1371/journal.pone.0185613.
12. Hariharan, M.; Polat, K.; Sindhu, R. A New Hybrid Intelligent System for Accurate Detection of Parkinson's Disease. *Comput Methods Programs Biomed* **2014**, *113* (3), 904–913. https://doi.org/10.1016/j.cmpb.2014.01.004..
13. Sakar, C. O.; Serbes, G.; Gunduz, A.; Tunc, H. C.; Nizam, H.; Sakar, B. E.; Tutuncu, M.; Aydin, T.; Isenkul, M. E.; Apaydin, H. A Comparative Analysis of Speech Signal Processing Algorithms for Parkinson's Disease Classification and the Use of the Tunable Q-Factor Wavelet Transform. *Applied Soft Computing* **2019**, *74*, 255–263. https://doi.org/10.1016/j.asoc.2018.10.022.
14. Solana-Lavalle, G.; Galán-Hernández, J.-C.; Rosas-Romero, R. Automatic Parkinson Disease Detection at Early Stages as a Pre-Diagnosis Tool by Using Classifiers and a Small Set of Vocal Features. *Biocybernetics and Biomedical Engineering* **2020**, *40* (1), 505–516. https://doi.org/10.1016/j.bbe.2020.01.003.
15. Sakar, C.O.; Serbes, G.; Gunduz, A.; Tunc, H.C.; Nizam, H.; Sa kar, B.E.; Tutuncu, M.; Aydin, T.; Isenkul, M.E.; and Apaydin, H. A comparative analysis of speech signal processing algorithms for Parkinsonâ€™s disease classification and the use of the tunable Q-factor wavelet transform. *Applied Soft Computing* **2017**, DOI: https://doi.org/10.1016/j.asoc.2018.10.022.
16. Murty, K. S. R.; Yegnanarayana, B. Combining Evidence from Residual Phase and MFCC Features for Speaker Recognition. *IEEE Signal Processing Letters* **2006**, *13* (1), 52–55. https://doi.org/10.1109/LSP.2005.860538.
17. Selesnick, I. W. Wavelet Transform With Tunable Q-Factor. *IEEE Transactions on Signal Processing* **2011**, *59* (8), 3560–3575. https://doi.org/10.1109/TSP.2011.2143711.
18. Al Ghayab, H. R.; Li, Y.; Siuly, S.; Abdulla, S. A Feature Extraction Technique Based on Tunable Q-Factor Wavelet Transform for Brain Signal Classification. *J Neurosci Methods* **2019**, *312*, 43–52. https://doi.org/10.1016/j.jneumeth.2018.11.014.
19. Amarappa, S.; Sathyanarayana, S. Data Classification Using Sup port Vector Machine (SVM), a Simplified Approach. Int. J. Elec tron. *Comput. Sci. Eng* **2014**, *3*, 435–445.
20. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; others. Scikit-Learn: Machine Learning in Python. t*he Journal of machine Learning research* **2011**, *12*, 2825–2830.
21. Louppe, G. Accelerating Random Forests in Scikit-Learn. **2014**.

https://orbi.uliege.be/bitstream/2268/171887/1/slides.pdf
22. Trappenberg, T. P. Machine Learning with Sklearn. *In Fundamentals of Machine Learning;* Oxford University Press, 2019; pp 38–65.
23. Janitza, S.; Hornung, R. On the Overestimation of Random Forest's out-of-Bag Error. *PLoS One* **2018**, *13* (8), e0201904. https://doi.org/10.1371/journal.pone.0201904.

Author

Aadya Bhat is a junior (Class of 2023) at Eastlake High School in Sammamish, Washington. She is particularly interested in neuroscience and hopes to pursue higher education in this field and healthcare/medicine. This paper was inspired by her experiences as a volunteer in a hospital's Neuro-Epilepsy and Neuro-Telemetry wards.

DOI: 10.36838/v4i6.1

IJHSR
ijhighschoolresearch.org
REVIEW ARTICLE

Contemporary Trends in MRI and PET Imaging for the Early Diagnosis of Alzheimer's Disease

Bianca Chan
Green Hope High School, 2500 Carpenter Upchurch Rd, Cary, North Carolina, 27519, USA; biancachan1025@gmail.com

ABSTRACT: Recent advancements in Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET) have allowed us to measure brain changes that occur during various stages of Alzheimer's Disease (AD). The utilization of these tools has arguably enabled us to effectively characterize important neural features of AD. Using Structural MRI, researchers have measured global and regional brain tissue loss and ventricular enlargement in AD Patients, evaluated by cognitive decline and AD clinical manifestations. Using Functional MRI and Perfusion MRI, researchers have detected abnormal network dysfunction and alterations in blood flow in the brains of AD patients, which occur along with neuropathological changes. Using PET, scientists examined reduced metabolic brain activity and glucose levels in certain brain cortices of AD patients. Although the exact cellular-level mechanisms behind many of these neural phenomena are still unclear, these imaging methods have allowed us to dive deeper into the neural correlates of AD and have brought the possibility of diagnosing AD at the pre-symptomatic stage closer to reality.
KEYWORDS: Alzheimer's Disease; neural biomarkers; medial temporal lobe atrophy; default mode network; hypoperfusion.

Introduction

Since the German physician, Dr. Alois Alzheimer, documented the existence of Alzheimer's Disease in 1906, much insight has been gained about this neurological condition. Alzheimer's Disease (AD), a prominent type of senile dementia, is a fatal neurodegenerative disorder most prevalent in individuals over the age of 60.[1] Although the etiology of AD is not fully understood, the two primary the histopathological hallmarks of this illness in its early stages are notably the deposition of extracellular β-amyloid plaques and intraneuronal tangles of hyperphosphorylated tau proteins across the brain, specifically within the medial temporal lobe.[2] Generally, β-amyloid plaques first aggregate in the basal cortex before spreading to other brain structures.[3] Significant tau neurofibrillary tangles have been found to correlate with subtle cognitive dysfunction in individuals without AD, hinting that it may be present during incipient AD or other pre-dementia stages, like mild cognitive impairment (MCI). These tangles first clump together in the transentorhinal area before spreading throughout the limbic area and in the association cortices.[4] Specifically, the spread of tau tangles defines the progression of the Braak stage of AD, which is made up of the so-called transentorhinal, limbic, and isocortical stages, respectively.[5] Deposition of these proteins induces gradual neuronal damage in individuals with AD, leading to AD's clinical symptoms. Moreover, protein deposition impedes neuronal communication before cell death, which also leads to memory problems, difficulties in word-finding, and the eventual, general cognitive decline and difficulty handling day-to-day tasks apparent in individuals suffering from AD.[6] Though these proteins are salient in the brains of individuals with AD, the current blood testing methods of these proteins are neither standardized nor specific or sensitive enough to confirm AD.[7] In fact, current diagnostic tests for AD-related proteins are unable to differentiate AD from other types of dementia. Medical imaging methods, such as PET imaging, are not only able to quantify the accumulation of amyloid-beta and tau pathology in the brain but can also precisely capture the neural changes associated with AD.[8] Novel developments in AD detection—MRI and PET neuroimaging of AD structural and functional measures—are the primary focus of this review.

Importantly, neurophysiological changes in individuals who eventually get diagnosed begin years before clinical symptoms manifest, a state called incipient AD. These changes — structural and functional — within the central nervous system are detectable using Magnetic Resonance Imaging (MRI). This non-invasive *in-vivo* approach has continued to contribute to our current scientific field of knowledge on AD. Moreover, it increasingly provides scientists hope for the prospect of using MRI to aid in detecting this illness before onset and inhibiting the further development of this disease through the rendering of therapeutic disease-modifying drugs. In this research paper, we will review the current knowledge around Structural MRI (sMRI), Functional MRI (fMRI), and Positron Emission Tomography (PET) correlates of AD while emphasizing current high-resolution MRI analytic techniques for detecting the presence of AD before an individual is clinically affected or pathologically diagnosed.

Structural MRI Neural Correlates of Alzheimer's Disease:

Structural Magnetic Resonance Imaging (sMRI), which provides white matter/gray matter contrast, has become a well-validated assessment of the rate and amount of tissue damage in characteristically AD-susceptible brain regions such as the medial temporal lobe. The primary use of sMRI for the detection of atrophy in AD uses automated methods to trace the change in thickness of specific regions of interest—for instance, the hippocampus or entorhinal cortex—over time. Among the primary regions of interest, it has become clear that hippocampal volume loss is a major correlate of AD, and damage to this area is a likely cause of AD symptoms,

DOI: 10.36838/v4i6.2

given that the hippocampus is an early site of tau deposition and MRI-based atrophic volume changes. Specifically, damage typically lies along the perforant or polysynaptic hippocampal pathway. However, it is important to note that the hippocampus receives its input from the entorhinal cortex, meaning that tau deposition typically develops first in the entorhinal region, and damage to the area leads to hippocampal dysfunction. As the hippocampus plays a major role in mediating memory formation and retrieval, memory loss often characterizes the first symptoms of AD and is thought to culminate from atrophy of the hippocampus and the nearby entorhinal cortex.[9]

Additionally, sMRI has been used to measure global brain atrophy and appraise increased ventricular volumes. Studies indicate that the topography of tissue loss aligns with advancing cognitive deficits seen in symptomatic AD individuals, both cross-sectionally and longitudinally.[10] Using sMRI, scientists can assess the various pathological and neurobiological dynamics that correspond to the structural changes in the brain across a range of stages in the progression of AD. In doing so, we will evaluate how well tissue volume loss detected by sMRI serves as a neurodegenerative marker of AD and its progression.

Regional atrophy in AD and MCI:

Brain atrophy in particular regions, particularly the medial temporal lobe, has been able to distinguish patients with AD from healthy individuals, and thus serves as an advantageous marker for the diagnosis of the disease. Whereas the rate of normal aging atrophy is only about 0.7% annually, the rate of disease-related atrophy is about 1.4-2.2%.[10] This means that the amount and degree of atrophy are much higher in AD than atrophy in healthy aging persons. Moreover, this rate increases approximately 0.3% annually in the years leading up to AD or Dementia diagnosis and it also accelerates as the brain transitions from normal cognitive performance to cognitive impairment, i.e., when MCI progresses to AD. Brain atrophy is also visible in individuals years before developing AD and before any manifestation of clinical symptoms.[11] This longitudinal study found that subjects who would later become AD patients revealed a reduction in left hippocampus volumes in the baseline scan compared to controls. At the end of the study, subjects who developed AD exhibited a hippocampal volume loss of more than 8% annually. In comparison, the hippocampal volume of the individuals who remained healthy did not display substantial differences from the controls at any stage throughout the study, with a mere 1% loss in hippocampal volume in the 3 years of the study. In corroboration, subjects who would later be diagnosed with AD showed a significant degradation in brain tissue volumes in several brain regions such as the medial temporal lobes, posterior cingulate/precuneus, and orbitofrontal cortex in the baseline scan.[12] There was also a reduction in gray matter (GM) density in the right medial temporal lobe and posterior cingulate in individuals who would later develop AD and in individuals who were destined to have Mild Cognitive Impairment (MCI) when compared to controls. MCI is also known as prodromal AD and represents a transitional, clinical phase between normal cognitive aging and AD (although not all MCI cases are subject to conversion to AD).[13]

There appears to be a hemispheric effect to these findings. Atrophy in the right hippocampus is greater and occurs earlier than the atrophy of the left hippocampus.[14] On the other hand, there is a pattern of a greater volume of the right hippocampus than the volume of the left hippocampus in normal aging individuals, a trend that is flipped through the progression of AD. This dynamical asymmetry change leads to the notion that there is often more right hippocampal volume loss than left in AD patients. Overall, the consensus is that brain atrophy in specific regions associated with AD pathology is present up to 10 years before the occurrence of clinical symptoms.[15] This suggests that hippocampal atrophy is identifiable and detectable in MRI scans in the early stages or years before symptomatic AD, although when these changes begin to occur is unclear. Perhaps these changes could be evaluated in correspondence with the first deposition of tau and β-amyloid protein, the two ingrained pathological occurrences of AD, to ascertain the timeline of brain atrophy.

Volume loss in other brain regions can also serve as a predictive measure of subsequent cognitive decline in AD. 6-month atrophy rates in medial and lateral temporal regions such as the left and right middle temporal gyri and left and right entorhinal cortices in healthy elderly, who would later develop AD, were prognostic of 2-year memory deterioration.[16] Carriers of the *APOE ε4* allele, a gene that possesses greater risk of late-onset AD to carriers, also have accelerated cortical thinning in the medial frontal and pericentral cortices, the area's most vulnerable and fragile to aging as well as regions associated with AD and amyloid aggregation preceding the diagnosis of AD.[17] Not only so, but these carriers who later develop AD demonstrated substantially reduced entorhinal cortex thickness in comparison with carriers who did not develop AD. Brain atrophy rates have also been shown to accelerate as an individual nears the diagnosis of AD, suggesting that brain tissue loss does not occur linearly. Regional and global-brain atrophy rates soared 3 to 5 years before being clinically affected with AD in a cohort of pre-symptomatic AD autosomal dominant mutation carriers, who are prone to develop early-onset familial AD.[18] All in all, tissue loss in the brain can foreshadow the diagnosis of hereditary AD as well as the cognitive impairment it brings in individuals with the *APOE ε4* gene.

Ventricular Enlargement in AD:

Ventricular enlargement is characteristic of all forms of gross neuronal damage and reflects the expansion of cerebrospinal fluid into regions formerly occupied by living neural tissue.[19] Naturally, ventricular enlargement is an explicit indicator of the advancement of Mild Cognitive Impairment and AD. For instance, total ventricular volumes in individuals who have AD and individuals who have MCI have been examined over 6 months, with a comparison to the total ventricle volumes of healthy subjects.[20] Though overlap existed in all groups in the follow-up, it was apparent that both AD and MCI subjects had greater mean baseline ventricular volumes than that healthy subjects and this difference grew over time. Compared to healthy individuals, subjects with AD also have

DOI: 10.36838/v4i6.2

a 4-fold greater enlargement over the same 6-month period. Ventricle size may also distinguish between AD and MCI; compared to subjects with MCI, Subjects with AD had a 60% greater ventricular enlargement over 6 months. Furthermore, it is interesting to note that lower baseline cognitive scores were associated with a higher baseline ventricular size and a higher rate of increased ventricle size over time. To illustrate the scale of these changes, the ventricles expand by about 16% annually in AD patients—a stark contrast with the annual expansion of 1.5 to 3.0% observed in healthy age-matched individuals.[21] Given the relationship between ventricle size and general neuronal damage, ventricle size may allow scientists to summarize the amount of neuronal damage in the brain and evaluate disease progression at regular intervals. Overall, it can be deduced that ventricular enlargement is an effective marker in the monitoring of AD. However, because ventricular enlargement is also present in MCI, it is not strong enough to serve as the sole diagnostic criteria for the transition between MCI and AD but should rather be considered alongside other measures.

Distinguishing AD and MCI Using Structural Changes:

Brain volume and cognitive dynamics in Mild Cognitive Impairment and AD have been explored and evaluated in comparison both with each other and with normal the elderly population. It is demonstrated that atrophy in both MCI and AD have strikingly similar patterns, although AD atrophy is naturally more severe than in MCI.[22] Though, differences besides the degree of atrophy may help distinguish MCI from AD. For instance, subjects with MCI show particularly less atrophy in medial temporal regions. Unsurprisingly, the extent to which the general pattern of regional atrophy in AD atrophy is mirrored in a given MCI subject is predictive of cognitive decline and progressive structural brain degradation.[23] In other words, MCI subjects whose brain tissue loss most resembles the typical regional AD atrophy pattern tend to undergo a considerably higher rate of decline and, indeed, conversion to AD. This suggests that quantitative measures of atrophy from structural MRIs are sensitive to AD neurodegeneration, and although atrophy itself is nonspecific to AD, the degree of atrophy may be a sensitive hallmark of AD histopathology and a predictor of cognitive decline severity once the diagnosis of AD is established by other means.

Both atrophies within the entorhinal cortex and hippocampus have been used to detect the conversion from a healthy state to MCI and from MCI to AD.[24] When distinguishing between healthy individuals and MCI patients, volume analysis using voxel-based methods suggests that the entorhinal cortex can serve as a better marker to distinguish between healthy controls and MCI patients, given that it deteriorates earlier than the hippocampus does in MCI. In other words, when AD patients are not included in the analysis, entorhinal cortex structural change is the most feasible classification to discriminate between healthy aging individuals and MCI individuals.[25] On the other hand, volumetric measurements of the hippocampus seem to be the most viable method for discriminating between AD patients from healthy controls and AD patients from MCI patients. This is in part due to its somewhat later deterioration and due to the inaccuracy introduced by the entorhinal cortex's small size, especially when it has heavily atrophied in the later stages of AD progression.[26] Further studies also corroborate these findings: one contends that the entorhinal cortex alone does not yield additional advantages beyond hippocampal atrophy for classification between AD and MCI,[27] while another, which first used hippocampal volume and later incorporated the entorhinal cortex into its model, found no increased effectiveness of classification between AD and MCI.[28]

Although global brain atrophy appears to be a robust biomarker to support the clinical diagnosis of AD, it is not a robust enough biomarker to diagnose AD at the MCI stage alone.[29] Instead, MCI patients are inclined to develop AD if they possess one of the following four biomarkers: a positive value in amyloid imaging, medial temporal atrophy, abnormality of cerebrospinal fluid markers (tau, β-amyloid 42, or phospho-tau), and temporoparietal cortical hypometabolism. Out of all the structural biomarkers, the evaluation of the degree of medial temporal lobe atrophy on a high-resolution T1-weighted MRI image using quantitative methods has been the single most pervasively used and most accurate for the diagnosis of AD. Although this may be true, medial lobe atrophy, when combined with other diagnostic markers such as abnormal CSF biomarkers, has resulted in even stronger predictive models.[30] In other words, this means that medial temporal lobe atrophy should be integrated with other AD pathological markers for diagnosis of AD during the MCI state, to produce optimal diagnostic and detection results.

Discussion

Structural imaging based on MRI is a pivotal tool for the detection, monitoring, and assessment of individuals with suspected AD before the onset of clinical symptoms. The atrophy of the medial temporal lobe—a marker sensitive to but not specific to AD–is a particularly interesting candidate for the early diagnosis of AD at the MCI stage and when the individual is still cognitively healthy, although the addition of other markers is likely producing even more reliable detection. Strikingly, the atrophy of the hippocampus can be detected up to 10 years before one is pathologically diagnosed with AD, and in comparison, with healthy aging individuals, atrophy in AD patients is significantly greater and faster. The rate of hippocampal and whole-brain atrophy, although not linear and consistent, are sensitive and robust markers for the progression of neurodegeneration. Increasingly, it seems that they can and should be measured alongside neuropsychology exams and tests such as the Mini-Mental State Exam, which is used to assess cognitive functions and abilities among the elderly, to support an MCI and/or AD diagnosis.

Ventricular enlargement is also a precise and suitable measure of neuropathological change associated with AD, as substantiated in multiple follow-up comparative studies. All in all, sMRI studies of these longitudinal changes of structural markers in AD susceptible patients offer relatively new diagnostic criteria for AD, potentially long before the traditional, pathological diagnosis. But despite the progress that has been made to understand the groundwork of AD neuropathology, there is still much to learn about the exact timeline

and first appearance of these structural markers. It remains unclear when to diagnose AD using these markers given that all of the reported findings are probabilistic. Moreover, though sMRI reflects changes in the brain, it does not detect AD histopathological hallmarks. Thus, the pathological cause of structural changes cannot be identified. Additionally, given that similar patterns of atrophy could also be found in other diseases, hence the nonspecific nature of AD structural measures, information extracted from sMRI should be integrated with other technologies such as PET imaging modalities of amyloid and tau to produce definitive evidence of AD.

Functional Neural Correlates of Alzheimer's Disease:

Positron emission tomography (PET) has long been used to assess the functional consequences of AD and is most useful in assessing the degree of metabolism remaining in the damaged brain of AD patients. Blood Oxygenation Level-Dependent (BOLD) fMRI (often used alongside PET) is another way to infer brain function, a more modern tool in the context of AD diagnosis. With the high spatial and reasonable temporal resolution it offers, BOLD fMRI may provide useful information about communication between discrete brain regions and, thus, the functional integrity of brain networks in support of cognitive function and their corresponding changes in AD. In recent years, both BOLD fMRI and PET have offered advanced approaches in the identification of the functional neural correlates of AD. Although BOLD fMRI and PET have been the most broadly used in the studying of AD, perfusion fMRI has also been used to track changes in cerebral blood flow in the AD brain which provides inferences about brain activity, similar to BOLD fMRI. Ultimately, the goal of these techniques in the context of AD is to define the functional alterations that occur before and during AD and how they relate to neuropathological changes.

Despite extensive investigation and examination, the underlying functional neural basis of AD remains mysterious. Here, this paper reviews the known functional correlates of AD, primarily changes in brain metabolism, cerebral blood flow, and the connectivity of the default mode network as detected using perfusion MRI, BOLD fMRI, and PET. Additionally, the paper corroborates these results to examine how these changes are associated with AD neuropathological biomarkers as measured using PET and how these biomarkers are exhibited across the clinical spectrum.

Default Mode Network Dysfunction:

In the past decade, the default mode network (DMN) has been defined as an intercorrelated network of brain regions that increase in activity and mutual connectivity when an individual is focused on his or her inner world.[31] The DMN is made up of the medial frontal cortex, the posterior cingulate cortex, and the bilateral angular gyrus/inferior parietal lobe but can extend to secondary regions such as the medial temporal lobe structures including the hippocampus.[32] Changes within the DMN have been commonly detected in AD patients and populations at risk for AD, which includes amnestic Mild Cognitive Impairment (aMCI) patients and even asymptomatic *APOE-ε4* carriers.[33] Within the brain structures implicated in this network, the inferior parietal lobes and the posterior cingulate cortex have appeared to be the most affected in AD.[34]

For instance, AD patients demonstrate a general deficiency of resting-state network activity within the DMN in comparison with the resting state networks of normal aging individuals.[35] Given that the AD participants as a whole had significantly lower hippocampal coactivity with the DMN, the study strongly suggested that the default mode network is closely involved in episodic memory processing. Additionally, to identify whether the default mode network of *APOE ε4* carriers is altered, one study recruited individuals with subjective memory complaints and compared those susceptible to AD due to an *APOE ε4* allele to those without that risk factor.[36] Indeed, the presence of *APOE ε4* correlates with default mode network changes, most present in frontal and posterior areas and the right hippocampus. All in all, this suggests that DMN disruption is robustly associated with AD with implications for hippocampal dysfunction, and may, in part, mediate the link between AD and *APOE ε4* carriers. However, future work should address the specific symptomatic correlates of the default mode network in AD and AD-susceptible individuals.[37]

Moreover, work has been conducted to better understand the link between network dysfunction and the cellular damage observed in AD. One study found that tau tends to distribute within the components of the neural networks ostensibly impacted by AD, such as in the default mode network.[38] It concluded that the aggregation of tau pathology in the brains of individuals with mild-to-moderate AD may disrupt several independent pathways and is closely associated with global cognitive dysfunction. Moreover, one also reports that as the Aβ burden increases regions outside of the DMN become less coactivated with DMN structures, namely the ventral medial PFC, right angular gyrus, and the left frontal gyri.[39] This demonstrates that the accumulation of tau and amyloid-beta protein, the two most prominent pathological hallmarks of AD, are likely involved in changes in local DMN function and its impact throughout the brain.

Despite the insight provided by recent investigations, the exact origin of default mode network changes in AD remains far from fully understood. Moreover, the current relationship between the AD structural changes and network changes is not well defined. While one study found that greater atrophy in AD generally is related to less brain interconnectivity,[40] another reported that the degree of functional connectivity changes is not associated with the severity of gray matter atrophy in the post-cingulate cortex, a central region of the DMN.[41] Thus, while structural changes must result in the functional alterations observed in AD, the precise mechanisms of this relationship are far from fully characterized amid early and contradictory findings.

AD Hypoperfusion:

In AD, perfusion or cerebral blood flow and volume have been measured with several different imaging modalities, including fMRI. Global and regional hypoperfusion, a reduction in the amount of the blood flow to the brain, which is evident in more than 60% of all AD patients, is an integral component in the understanding of the pathophysiology of this illness.

DOI: 10.36838/v4i6.2

Not only is it apparent throughout disease progression, but disrupted perfusion is also present during MCI as well as in cognitively healthy individuals with the *APOE ε4* allele at high risk for developing familial AD.[42] This phenomenon is associated with both structural and functional changes in the brain and may potentially serve as a biomarker for the identification of AD in its preclinical phase.

According to one study, individuals with AD demonstrate a dramatic global decrease (~40%) in blood flow compared to healthy controls.[43] In particular, the parietal lobe demonstrates decreased regional cerebral blood flow (CBF) in AD.[44] One study also unveiled particularly reduced CBF in the precuneus and the posterior cingulate in individuals with AD when compared with healthy individuals.[45] Other structures associated with decreased CBF include the temporo-occipital and parietal-occipital cortices, hippocampus, parahippocampus, bilateral inferior regions, and areas in the prefrontal cortex, such as the bilateral superior and middle frontal gyri.[46] Moreover, in regions with gray matter loss or atrophy such as the right inferior parietal lobe and posterior cingulate, individuals with AD have persistently shown diminished CBF.[47]

The pattern of altered perfusion in AD and MCI patients has also been compared and contrasted using pulsed arterial spin labeling (PASL), a specialized technique designed to evaluate brain perfusion. In comparison to healthy subjects, investigations of perfusion in MCI individuals have demonstrated a decrease in CBF in the posterior circulate extending to the medial precuneus and right inferior lobe.[48] One study compared AD and MCI individuals with healthy controls and found decreased CBF in both the AD and MCI group.[49] However, the AD individuals as a whole demonstrated greater hypoperfusion than the MCI group, particularly in regions such as the bilateral precuneus and the PCC. The consistent observation of PCC changes is interesting given its significant role in the DMN while consistently decreased perfusion within the precuneus may be relevant to the symptoms of AD, given the purported role of this region in memory as well cognitive functioning more broadly. Although extensive research has authenticated hypoperfusion to be a robust hallmark of AD, it is still unclear how this disruption of blood flow contributes to clinical manifestation across the disease spectrum. Its direct co-occurrence with other neuropathological mechanisms of AD such as amyloid deposits has rendered it difficult to determine its specific causal impact on AD.

AD hypometabolism:

One of the predominant functional hallmarks of AD is glucose hypometabolism, detected using PET. Particularly, there is abnormally low posterior and whole-brain metabolism in patients with severe AD, and there is a gradual decline in these levels over time. The progressive reduction in glucose levels has also been found to align with cognitive performance, corroborating that glucose level decline in the brain is a reliable and validated marker of the AD stage and status.[50] As with the other correlates mentioned in this review, the reduction of cerebral glucose metabolism often occurs before clinical symptoms manifest and continues as symptoms progress. It is also more severe in populations at risk for AD and populations with AD.

For instance, one study found that individuals with *APOE ε4* allele who would later develop AD had a substantial reduction in the rate of glucose metabolism, in comparison with the controls, in the posterior cingulate, parietal, temporal, and prefrontal regions, sites of early deposition of neuropathological lesions.[51] Similarly, another found that cognitively healthy carriers of the APOE ε4 allele have an unconventionally low cerebral metabolic rate for glucose in the same regions where hypometabolism is most notable in AD patients.[52]

One presumed explanation for reductions in glucose levels in AD patients is neuronal loss or brain atrophy, which is a chief hallmark of the disease. However, a marked decrease in glucose metabolism, which provides the fuel for brain physiological functions and activity, was found in presymptomatic incipient AD individuals in the absence of structural brain atrophy, suggesting brain tissue damage or cell loss is not the only contributor to these low glucose levels. Because regional decreased glucose metabolism appears to not be a result of cell atrophy, researchers have been investigating other neuropathological causes. For instance, one study examined the effects of amyloid precursor protein (APP) on hypometabolism before death.[53] It detected that heme deficiency on APP could catalyze heme metabolism in AD. Additionally, another outlined that disturbances in the insulin-lipid-glucose axis may lead to hypometabolism.[54] It proposed that the downstream effects that transpire as a result of hypometabolism may in part be due to reduced production of acetyl-coenzyme A (CoA) and ATP, both of which are involved in biochemical reactions in lipid, protein, and carbohydrate metabolism and are crucial for synaptic plasticity and activity. Although these endeavors have brought the constituents of the etiology of AD glucose hypometabolism into light, the precise mechanism that accounts for this neural correlate is still widely unknown.

Discussion

BOLD fMRI, perfusion fMRI, and PET have identified several consistent functional neural underpinnings of AD. Foremost, the default mode network activity has been consistently found to be disrupted and impeded in AD individuals and individuals at high risk for AD, in comparison with cognitively healthy individuals. Even more so, its altered interactions with the hippocampus appear particularly relevant in episodic memory processing, explaining the memory dysfunction often seen in AD patients. Pathological features of AD have also been examined and measured about their impacts on the default mode network. Specifically, the aggregation of proteins such as tau and amyloid-beta has been shown to also hinder the default mode network pathway while the *APOE ε4* allele is associated with DMN disruption in non-AD subjects. These findings secure DMN disruption as a primary correlate of AD, a candidate for early diagnosis, and a potentially crucial mechanistic underpinning of AD symptoms.

Perfusion fMRI has illuminated the role of cerebral blood flow in AD. Disrupted perfusion is not only evident throughout disease manifestation, but it is also demonstrated during

the preclinical phase of AD (i.e., MCI) as well as in cognitively healthy persons at high risk for developing AD due to family history or genetic factors. Although evidence is limited and the relationship between perfusion and the symptoms of AD remains unclear, changes in perfusion have the potential to act as a diagnostic marker of this illness.

Additionally, increasing evidence from PET denotes depletion of brain glucose metabolism as an early and progressive incident in AD. This reduction in metabolic rate is observed particularly within the areas of tau and amyloid-beta deposition, in both symptomatic AD patients and presymptomatic individuals with the *APOE ε4* gene. More so, this decrease in glucose levels is in parallel with progressive, gradual cognitive decline, offering hypometabolism as yet another potent marker in the tracking of this disease. In summation, these cardinal neural correlates constitute the primary functional changes that are present in the brain in prodromal AD and throughout its stages of progression.

Conclusion

Although neuroimaging has provided many novel correlates of AD, it remains unclear how the neuropathological, structural, and functional correlates of this neurologic disorder relate to one another.

Foremost, despite the well-known occurrence of brain atrophy and the presence of abnormal protein build-up in AD, the exact contribution of proteins clumping to tissue volume loss over time has not been thoroughly examined. Perhaps protein footprintings, a technique of biochemical analysis that maps protein structure, could serve to integrate tissue volume loss with protein aggregation in the brain.[55] Furthermore, detecting the specific proteins that give rise to atrophy through blood tests could also potentially help contribute to an even more reliable diagnosis of this disease. Predictive models built on these combined measures would sit at the crossroad between structural changes and the cellular-level changes that underlie them, perhaps aiding in early diagnosis and treatment options.

Secondly, a well-defined correspondence between how structural alterations of AD—e.g., ventricular enlargement and brain atrophy—account for network deficiency and connectivity impediment is still lacking. Uncovering the association between the AD structural and functional correlates is critical as this could provide an explanation for why changes in network activity are present in AD. One potential method is a relatively novel approach called "parallel group ICA+ICA", which links functional networks with structural variation by assimilating fMRI information with information from other modalities such as sMRI.[56] This methodology could outline the relationship between these two aspects of AD; this is of crucial importance as connecting them could yield unprecedented insight into the pathogenesis of this illness.

Together, structural changes and functional changes should be integrated with the manifestation of clinical symptoms, and the severity of such symptoms. This is paramount in the tracking and determination of an affected individual's disease stage. By doing so, scientists can more reliably account for neurophysiology symptoms or behavioral changes in correspondence evidence of structural or functional occurrences in sMRI and fMRI. For example, by bridging the gap between neural changes and clinical symptoms, scientists could answer the questions: "how is the severity of memory impairment directly attributed to the degree of hippocampal atrophy" or "How does low default mode network activity exhibit as a symptom in AD individuals?" One feasible approach could be voxel-based symptom mapping, which has been used to investigate certain brain regions where the occurrence of exteroceptive and proprioceptive somatosensory deficits in strokes is prevalent as well as examine large-scale changes in resting-state thalamocortical functional connectivity in parallel with mental impairments in Schizophrenia.[57,58] Voxel-based lesion-symptom mapping could be conducted in similar manners in AD patients to bridge the gap between neural changes and the symptoms associated with AD.

As the field progresses, novel answers to current limitations could provide for more optimal diagnosis, examination, and comprehension of this devastating neurodegenerative disease.

Acknowledgements

Special thank you to my mentor, Joshua Cain, and other editors both in and outside of the Lumiere Research Program for guiding me through this research project. I also thank my parents for supporting me through this project.

References

1. Frisoni GB; Fox NC; Jack CR; Scheltens P; Thompson PM; The clinical use of structural MRI in Alzheimer's disease. https://pubmed.ncbi.nlm.nih.gov/20139996/
2. Marks, S. M.; Lockhart, S. N.; Baker, S. L.; Jagust, W. J. Tau, and β-amyloid are associated with medial temporal lobe structure, func tion, and memory encoding in normal aging. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5373113 .
3. Palmqvist, S.; Insel, P. S.; Zetterberg, H.; Blennow, K.; Brix, B.; Stomrud, E.; Alzheimer's Disease Neuroimaging Initiative; Swed ish BioFINDER study; Mattsson, N.; Hansson, O. Accurate risk estimation of β-amyloid positivity to identify prodromal Alzhei mer's disease: Cross-validation study of practical algorithms. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6374284/
4. Adams, J. N.; Maass, A.; Harrison, T. M.; Baker, S. L.; Jagust, W. J. Cortical tau deposition follows patterns of entorhinal functional connectivity in aging. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6764824/
5. Dallaire-Théroux, C.; Beheshti, I.; Potvin, O.; Dieumegarde, L.; Saikali, S.; Duchesne, S.; Alzheimer's Disease Neuroimaging Ini tiative. Braak neurofibrillary tangle staging prediction from in vivo MRI metrics. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6731211/
6. Y; N. T. T. H. K. Neuronal cell death in Alzheimer's disease and a neuroprotective factor, Humanin. https://pubmed.ncbi.nlm.nih.gov/18615127/
7. Blood test methods may predict Alzheimer's protein deposits in the Brain. https://www.nia.nih.gov/news/blood-test-method-may-predict-alzheimers-protein-deposits-brain
8. Johnson, K. A.; Fox, N. C.; Sperling, R. A.; Klunk, W. E. Brain Imaging in Alzheimer's disease. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3312396/
9. Esther Heerema, M. S. W. How you can prevent the hippo campus from shrinking. https://www.verywellhealth.com/what-is-the-hippocampus-p2-98810
10. Frisoni GB; Fox NC; Jack CR; Scheltens P ;Thompson PM; The

DOI: 10.36838/v4i6.2

clinical use of structural MRI in Alzheimer's disease. https://pubmed.ncbi.nlm.nih.gov/20139996/
11.Fox NC; Warrington EK ;Freeborough P A;Hartikainen P;Kennedy AM;Steven s JM;Rossor MN; Presymptomatic hip pocampal atrophy in Alzheimer's disease. A longitudinal MRI study. https://pubmed.ncbi.nlm.nih.gov/9010004/.
12.Tondelli, M.; Wilcock, G. K.; Nichelli, P.; Jager, C. A. D.; Jenkinson, M.; Zamboni, G. Structural MRI changes detectable up to ten years before Clinical Alzheimer's disease. https://www.sciencedirect.com/science/article/pii/S0197458011002016
13.Angelucci F;Spalletta G;di Iulio F;Ciaramella A;Salani F;Colantoni L;Varsi AE;Gianni W;Sancesario G;Caltagirone C;Bossù P; Alzheimer's disease (AD) and mild cognitive impairment (MCI) patients are characterized by increased BDNF serum levels. https://pubmed.ncbi.nlm.nih.gov/20205668/
14.Barnes, J.; Bartlett, J. W.; van de Pol, L. A.; Loy, C. T.; Scahill, R. I.; Frost, C.; Thompson, P.; Fox, N. C. A meta-analysis of hippocampal atrophy rates in Alzheimer's disease. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2773132/.
15.Tondelli, M.; Wilcock, G. K.; Nichelli, P.; Jager, C. A. D.; Jenkinson, M.; Zamboni, G. Structural MRI changes detectable up to ten years before Clinical Alzheimer's disease. https://www.sciencedirect.com/science/article/pii/S0197458011002016
16.Murphy EA;Holland D;Donohue M;McEvoy LK;Hagler DJ;Dale AM;Brewer JB; ; Six-month atrophy in MTL structures is associated with subsequent memory decline in elderly controls. https://pubmed.ncbi.nlm.nih.gov/20633660/
17.Liu, C.-C.; Liu, C.-C.; Kanekiyo, T.; Xu, H.; Bu, G. Apolipoprotein E and Alzheimer's disease: Risk, mechanisms and therapy. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3726719/
18.Kinnunen, K. M.; Cash, D. M.; Poole, T.; Frost, C.; Benzinger, T. L. S.; Ahsan, R. L.; Leung, K. K.; Cardoso, M. J.; Modat, M.; Malone, I. B.; Morris, J. C.; Bateman, R. J.; Marcus, D. S.; Goate, A.; Salloway, S. P.; Correia, S.; Sperling, R. A.; Chhatwal, J. P.; Mayeux, R. P.; Brickman, A. M.; Martins, R. N.; Farlow, M. R.; Ghetti, B.; Saykin, A. J.; Jack, C. R.; Schofield, P. R.; McDade, E.; Weiner, M. W.; Ringman, J. M.; Thompson, P. M.; Masters, C. L.; Rowe, C. C.; Rossor, M. N.; Ourselin, S.; Fox, N. C.; Dominantly Inherited Alzheimer Network (DIAN). Presymptomatic atrophy in autosomal dominant Alzheimer's disease: A serial magnetic resonance imaging study. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5751893/
19.Apostolova, L. G.; Green, A. E.; Babakchanian, S.; Hwang, K. S.; Chou, Y.-Y.; Toga, A. W.; Thompson, P. M. Hippocampal atrophy and ventricular enlargement in normal aging, mild cognitive impairment (MCI), and Alzheimer disease. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3286134/
20.Nestor, S. M.; Rupsingh, R.; Borrie, M.; Smith, M.; Accomazzi, V.; Wells, J. L.; Fogarty, J.; Bartha, R.; Alzheimer's Disease Neuroimaging Initiative. Ventricular enlargement as a possible measure of Alzheimer's disease progression validated using the Alzheimer's disease neuroimaging initiative database https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2724905/#:~:text=Ventricular%20enlargement%20may%20be%20an,progression%20for%20multi%2Dcentre%20studies.
21.Frisoni GB;Fox NC;Jack CR;Scheltens P;Thompson PM; The clinical use of structural MRI in Alzheimer's disease. https://pubmed.ncbi.nlm.nih.gov/20139996/
22.McEvoy, L. K.; Brewer, J. B. Quantitative structural MRI for early detection of Alzheimer's disease. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3182103/
23.McEvoy LK;Holland D;Hagler DJ;Fennema-Notestine C;Brewer JB;Dale AM; ; Mild cognitive impairment: Baseline and longitudinal structural MR imaging measures improve predictive prognosis. https://pubmed.ncbi.nlm.nih.gov/21471273/
24.Leandrou S;Petroudi S;Kyriacou PA;Reyes-Aldasoro CC;Pattichis CS; Quantitative MRI brain studies in mild cognitive impairment and Alzheimer's disease: A methodological review. https://pubmed.ncbi.nlm.nih.gov/29994606/
25.E;, B. H. B. Neuropathological stageing of alzheimer-related changes. https://pubmed.ncbi.nlm.nih.gov/1759558/
26.Leandrou S;Petroudi S;Kyriacou PA;Reyes-Aldasoro CC;Pattichis CS; Quantitative MRI brain studies in mild cognitive impairment and Alzheimer's disease: A methodological review.
27.Du, A. T.; Schuff, N.; Zhu, X. P.; Jagust, W. J.; Miller, B. L.; Reed, B. R.; Kramer, J. H.; Mungas, D.; Yaffe, K.; Chui, H. C.; Weiner, M. W. Atrophy rates of the entorhinal cortex in AD and normal aging. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1851672/
28.Pennanen, C.; Kivipelto, M.; Tuomainen, S.; Hartikainen, P.; Hänninen, T.; Laakso, M. P.; Hallikainen, M.; Vanhanen, M.; Nissinen, A.; Helkala, E.-L.; Vainio, P.; Vanninen, R.; Partanen, K.; Soininen, H. Hippocampus and entorhinal cortex in mild cognitive impairment and early AD. https://www.sciencedirect.com/science/article/abs/pii/S0197458003000848?via%3Dihub
29.Frisoni GB;Fox NC;Jack CR;Scheltens P;Thompson PM; The clinical use of structural MRI in Alzheimer's disease. https://pubmed.ncbi.nlm.nih.gov/20139996/
30.Leandrou S;Petroudi S;Kyriacou PA;Reyes-Aldasoro CC;Pattichis CS; Quantitative MRI brain studies in mild cognitive impairment and Alzheimer's disease: A methodological review. https://pubmed.ncbi.nlm.nih.gov/29994606/
31.Peter Pressman, M. D. The role of an fmri in monitoring brain activity. https://www.verywellhealth.com/what-is-the-default-mode-network-2488818
32.Mars, R. B.; Neubert, F.-X.; Noonan, M. P.; Sallet, J.; Toni, I.; Rushworth, M. F. S. On the relationship between the "Default Mode Network" and the "Social Brain". https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3380415/
33.Mevel, K.; Chételat, G.; Eustache, F.; Desgranges, B. The default mode network in healthy aging and Alzheimer's disease. https://www.hindawi.com/journals/ijad/2011/535816/
34.Palmqvist S;Schöll M;Strandberg O;Mattsson N;Stomrud E;Zetterberg H;Blennow K;Landau S;Jagust W;Hansson O; Earliest accumulation of β-amyloid occurs within the default-mode network and concurrently affects brain connectivity. https://pubmed.ncbi.nlm.nih.gov/29089479/
35.Greicius, M. D.; Srivastava, G.; Reiss, A. L.; Menon, V. Default-mode network activity distinguishes Alzheimer's disease from healthy aging: Evidence from functional MRI. https://www.pnas.org/content/101/13/4637
36.Chiesa, P. A.; Cavedo, E.; Vergallo, A.; Lista, S.; Potier, M.-C.; Habert, M.-O.; Dubois, B.; Schotten, M. T. de; Hampel, H. Differential default mode network trajectories in asymptomatic individuals at risk for Alzheimer's disease. https://www.sciencedirect.com/science/article/pii/S1552526019300779
37.Greicius, M. D.; Srivastava, G.; Reiss, A. L.; Menon, V. Default-mode network activity distinguishes Alzheimer's disease from healthy aging: Evidence from functional MRI. https://www.pnas.org/content/101/13/4637
38.Hoenig MC;Bischof GN;Seemiller J;Hammes J;Kukolja J;Onur ÖA;Jessen F;Fliessbach K;Neumaier B;Fink GR;van Eimeren T;Drzezga A; Networks of tau distribution in alzheimer's disease. https://pubmed.ncbi.nlm.nih.gov/29315361/
39.Li, M.; Zheng, G.; Zheng, Y.; Xiong, Z.; Xia, R.; Zhou, W.; Wang, Q.; Liang, S.; Tao, J.; Chen, L. Alterations in resting-state func

tional connectivity of the default mode network in amnestic mild cognitive impairment: An fmri study - BMC Medical Imaging. https://bmcmedimaging.biomedcentral.com/articles/10.1186/s12880-017-0221-9
40.Hoffstaedter F;Grefkes C;Roski C;Caspers S;Zilles K;Eickhoff SB; Age-related decrease of functional connectivity additional to gray matter atrophy in a network for movement initiation. https://pubmed.ncbi.nlm.nih.gov/24399178/
41.Gili T;Cercignani M;Serra L;Perri R;Giove F;Maraviglia B;Caltagirone C;Bozzali M; Regional brain atrophy and function al disconnection across Alzheimer's disease evolution. https://pubmed.ncbi.nlm.nih.gov/20639384/
42.Austin BP;Nair VA;Meier TB;Xu G;Rowley HA;Carlsson CM;Johnson SC;Prabhakaran V; Effects of hypoperfusion in Alz heimer's disease. https://pubmed.ncbi.nlm.nih.gov/21971457/
43.Asllani I;Habeck C;Scarmeas N;Borogovac A;Brown TR;Stern Y; Multivariate and univariate analysis of continuous arterial spin la beling perfusion MRI in Alzheimer's disease. https://pubmed.ncbi.nlm.nih.gov/17960142/
44.Pávics L;Grünwald F;Reichmann K;Horn R;Kitschenberg A;Hartmann A;Menzel C;Schomburg AG;Overbeck B;Cser nay L;Biersack HJ; Regional cerebral blood flow single-photon emission tomography with 99mtc-HMPAO and the acetazolamide test in the evaluation of vascular and alzheimer's dementia. https://pubmed.ncbi.nlm.nih.gov/10079314/
45.PMC, E. https://europepmc.org/article/MED/20413865
46.Austin BP;Nair VA;Meier TB;Xu G;Rowley HA;Carlsson CM;Johnson SC;Prabhakaran V; Effects of hypoperfusion in Alz heimer's disease. https://pubmed.ncbi.nlm.nih.gov/21971457/
47.Dai W;Lopez OL;Carmichael OT;Becker JT;Kuller LH;Gach HM; Mild cognitive impairment and Alzheimer's disease: Patterns of altered cerebral blood flow at Mr Imaging. https://pubmed.ncbi.nlm.nih.gov/19164119/
48.Austin BP;Nair VA;Meier TB;Xu G;Rowley HA;Carlsson CM;Johnson SC;Prabhakaran V; Effects of hypoperfusion in Alz heimer's disease. https://pubmed.ncbi.nlm.nih.gov/21971457/
49.Johnson NA;Jahng GH;Weiner MW;Miller BL;Chui HC;Jagust WJ;Gorno-Tempini ML;Schuff N; Pattern of cerebral hypoperf usion in alzheimer disease and mild cognitive impairment mea sured with arterial spin-labeling MR imaging: Initial experience. https://pubmed.ncbi.nlm.nih.gov/15734937/
50.Small GW;Ercoli LM;Silverman DH;Huang SC;Komo S;Bookheimer SY;Lavretsky H;Miller K;Siddarth P;Rasgon NL;Mazziotta JC;Saxena S;Wu HM;Mega MS;Cummings JL;Saunders AM;Pericak-Vance MA;Roses AD;Barrio JR;Phelps ME; Cerebral metabolic and cognitive decline in persons at ge netic risk for alzheimer's disease. https://pubmed.ncbi.nlm.nih.gov/10811879/
51.Reiman, E. M.; Al., E.; Author Affiliations From the Positron Emission Tomography Center; Campion, E. W.; Others, Y. M. a nd; Others, B. T. L. and; Others, J. W. and; Drain, P. K.; E. B. Walter and Others; Others, N. A. and. Preclinical evidence of Alzheimer's disease in persons homozygous for the ε4 allele for apo lipoprotein E: Nejm. https://www.nejm.org/doi/full/10.1056/nejm199603213341202
52.Reiman, E. M.; Caselli, R. J.; Chen, K.; Alexander, G. E.; Ban dy, D.; Frost, J. Declining brain activity in cognitively normal apo lipoprotein E epsilon 4 heterozygotes: A foundation for using positron emission tomography to efficiently test treatments to pre vent Alzheimer's disease. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC30654/
53.O'Brien, R. J.; Wong, P. C. Amyloid precursor protein processing and Alzheimer's disease. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3174086/
54.S;, H. Causes and consequences of disturbances of cerebral glucose metabolism in sporadic Alzheimer's disease: Therapeutic implica tions. https://pubmed.ncbi.nlm.nih.gov/14977212/
55.Bamberger C;Pankow S;Martínez-Bartolomé S;Ma M;Diedrich J;Rissman RA;Yates JR; Protein footprinting via covalent protein painting reveals structural changes of the proteome in Alzheimer's disease. https://pubmed.ncbi.nlm.nih.gov/33872013/
56.Qi S;Sui J;Chen J;Liu J;Jiang R;Silva R;Iraji A;Damaraju E;Salman M;Lin D;Fu Z;Zhi D;Turner JA;Bustillo J;Ford JM;Mathalon DH;Voyvodic J;McEwen S;Preda A;Belger A;Potkin SG;Mueller BA;Adali T;Calhoun VD; Parallel Group ICA+ICA: Joint estimation of linked Functional Network Variability and structural covariation with application to schizophrenia. https://pubmed.ncbi.nlm.nih.gov/31099151/
57.Meyer S;Kessner SS;Cheng B;Bönstrup M;Schulz R;Hummel FC;De Bruyn N;Peeters A;Van Pesch V;Duprez T;Sunaert S;Schrooten M;Feys H;Gerloff C;Thomalla G;Thijs V;Verheyden G; Voxel-based lesion-symptom mapping of stroke lesions un derlying somatosensory deficits. https://pubmed.ncbi.nlm.nih.gov/26900565/
58.Cheng, W.; Palaniyappan, L.; Li, M.; Kendrick, K. M.; Zhang, J.; Luo, Q.; Liu, Z.; Yu, R.; Deng, W.; Wang, Q.; Ma, X.; Guo, W.; Francis, S.; Liddle, P.; Mayer, A. R.; Schumann, G.; Li, T.; Feng, J. Voxel-based, brain-wide association study of aberrant functional connectivity in schizophrenia implicates thalamocortical circuitry. https://www.nature.com/articles/

Author

Bianca Chan is a sophomore in Green Hope High School. She has a strong passion for neuroscience. She decided to conduct Alzheimer's Disease research to shed light on this devastating neurodegenerative disease. Her work has been recognized by national scientific organizations and she aspires to study neuroscience in the future.

DOI: 10.36838/v4i6.2

IJHSR
ijhighschoolresearch.org
■ RESEARCH ARTICLE

Using Market Indices to Optimize Company Supply Chain

Charley L. Wan
Central Bucks East High School, 2804 Holicong Rd, Doylestown, PA 18902, USA; charleywan2004@gmail.com

ABSTRACT: The Supply chain triangle – Service, Cost, and Cash – is the core of any business. Providing the best service for customers, while reducing the supply chain cost and maximizing cash on hand are the three pillars of business operation. Businesses are constantly balancing among these three pillars of operation. Not being able to balance the three pillars of operation leads to the risk of developing bad service, high cost, or low cash.

Business planning is the act of using forecast market conditions to prepare future business operations and mitigate risks related to service, cash, and cost. A recent example is the COVID-19 pandemic's impact on businesses worldwide. Rapid changes in global demand and supply and government fiscal and monetary policies led to high inflation in prices. Now, more than ever, companies must incorporate market conditions into their planning to balance their supply chain triangles.

Risk pooling is a statistical concept that suggests that supply chain variability is reduced if one can aggregate variability across locations, products, or time. Modern Portfolio Theory (MPT) seeks to minimize non-systemic risk via a pool of diversified portfolios. Although MPT is frequently used in investment and finance, it is not found in supply chain management literature. This paper formulates a business supply chain plan as a portfolio optimization problem, which maximizes savings of inventory portfolios subject to constraints on inventory turns (service demand/inventory). This approach can further be used to achieve flexible operations by optimizing inventories for high-volume, low-variability materials to take advantage of economies of scale and/or optimizing inventories for low-volume, high-variability materials to pool the risks.

KEYWORDS: Supply Chain Triangle (Service, Cost, Cash), Supply and Demand, Risk Pooling, Market Indices, Modern Portfolio Theory (MPT), COVID-19 Pandemic, flexible operation.

■ Introduction

As an analogy, businesses are like ships floating in a vast ocean of the global economy (see Figure 1). In this ocean, it contains storms, such as inflation/deflation and politics/trade, that disrupt business operations. To fight against adverse market conditions, businesses need to control their "motors" (i.e., operation) and "steering" (i.e., planning) in order to navigate to success.

Figure 1: An analogy of a business in the global economy.

A core concept in economics is supply and demand.[1] On a macro level, two key situations have been impacting the current state and development of the world economy. Firstly, climate change has effects on different parts of the world including unexpected natural disasters and abnormal weather patterns. Pertaining to businesses, this causes plants to shut down and transportation disruptions.[2] Secondly, in the recent years of 2020-2021, the COVID-19 pandemic has caused people to lose jobs, due to quarantines worldwide.[3] Pertaining to businesses, consumer demand has been at an all-time low which then caused producers to bring supply down to very low levels as well. Although the markets are beginning to recover from in demand, supply cannot catch up, causing an imbalance in supply and demand and rapid rising in inflation4 (see Figures 2 and 3).

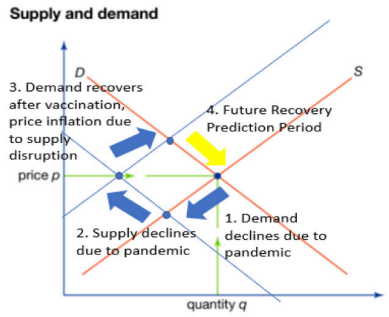


Figure 2: Supply and demand dynamics due to COVID-19.

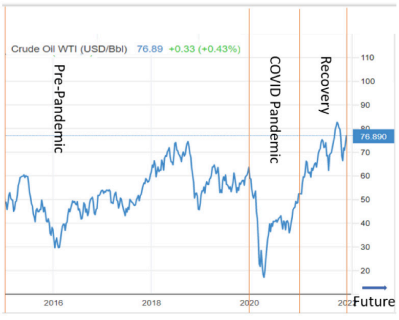


Figure 3: Crude oil WTI price volatility due to COVID-19[5]

DOI: 10.36838/v4i6.3

Although situations such as climate change and the COVID-19 pandemic represent systemic risks that businesses may not be able to control, businesses do have control over non-systemic risks by diversifying their portfolio. *Risk pooling is a statistical concept that suggests that supply chain variability is reduced if one can aggregate variability across locations, products, or time.*[6] *Modern Portfolio Theory (MPT),*[7,8] *which seeks to minimize non-systemic risk via a pool of diversified portfolio, can be applied to supply chain to find an optimal balance among services, cash, and cost.*

Traditionally, sales (the service side) will want to do anything to gain and maintain customers, which may force an increase in inventory and/or buying of materials even when costs are high. This may lead to a shortage of cash for a company. Finance (the cash side) will want to keep the inventory low and cash on hand as high as possible to keep the business afloat. This may limit sales in that they will not have products immediately available when customers need them. Procurement (the cost side) will want to base their decisions on predictions of market risk. When forecasts predict procurement costs will rise, they will want to buy early. Manufacturing and Logistics (the cost side as well) will both want to maximize operation efficiency and capacity utilization to process large batches/loads. This may conflict with sales when market demands are volatile, and conflict with procurement when market supplies are volatile[9] (See Figure 4).

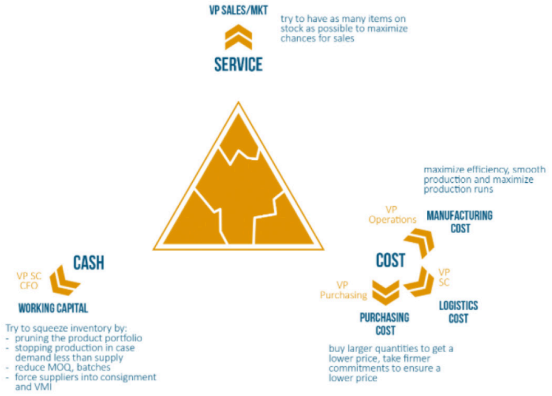


Figure 4: Breakdown of Supply Chain Triangle

This paper focuses specifically on the cash aspect (inventory) of the supply chain triangle and formulates the balance of the supply chain triangle as a portfolio optimization problem, which maximizes the savings of inventory portfolio subject to constraints on inventory turns (service demand/inventory). This approach can further incorporate the flexible operation rules by optimizing inventories for high volume-low, variability materials to take advantage of economies of scale at the secondary plants/warehouses and optimizing inventories for low-volume, high-variability materials to pool the risks at the primary plants/warehouses.

Methods

In this paper, PortfolioVisualizer.com[10] was used to solve portfolio optimization problems. WTI, PPI, CPI historical, and forecast data were obtained from TradingEconomics. WebDigitizer.[11]

Assuming the three materials have equal service demands and inventory turn 5~6, and have prices linearly correlated to Energy Market crude oil WTI, Producer Market PPI, and Consumer Market CPI, respectively (see Figures 5, 6).

Figure 5: Supply Chain

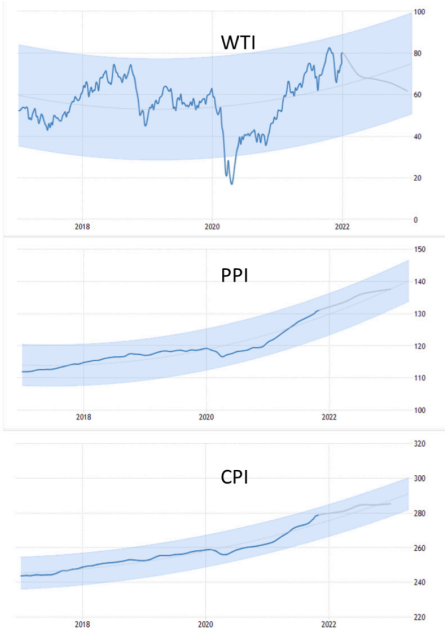


Figure 6: History and forecast time series of WTI, PPI, CPI

Results and Discussion

The objective of the proposed optimization is, like any investment portfolio optimization, to assemble a group of assets so that the portfolio has the highest return. In the context of the inventory portfolio, the objective is to stock a portfolio of materials as inventory so that it can provide maximum cost savings. Moreover, service and cash requirements are enforced by the constraint on inventory turn (service demand/inventory), (see Figure 7). Note, in the real world, there are many more constraints such as production/shipment constraints, product demand volume, variability constraints, etc.

Figure 7: A portfolio optimization approach for Supply Chain Planning

DOI: 10.36838/v4i6.3

Mathematically, the supply chain inventory portfolio optimization is formulated as follows:

$$maximize_{\alpha_1,\ldots,\alpha_n} = (\alpha_1 r_1 + \cdots + \alpha_n r_n)$$

Is subject to

$$\alpha_1 + \cdots + \alpha_n = 1$$

$$\alpha_k^{lower} \le \alpha_k \le \alpha_k^{upper}, for\ i = 1, \ldots, n$$

where r_1,…,r_n are returns of materials in the inventory portfolio, α_1,…,α_n are percentages of materials in the inventory portfolio, and

$$\alpha_k^{lower} = \left(\frac{demand_k}{demand_{total}}\right)\left(\frac{turn^{lower}}{turn^{upper}}\right)$$

$$\alpha_k^{upper} = \left(\frac{demand_k}{demand_{total}}\right)\left(\frac{turn^{upper}}{turn^{lower}}\right)$$

are lower and upper bounds of material k percentage in the inventory portfolio, which is material demand percentage multiplied by inventory turn lower-to-upper ratio and upper-to-lower ratio, respectively.

Table 1 summarizes the savings by using inventory portfolio optimization vs using equal weighted inventory. There is a range of inventory cost savings from 0.3% to 3%.

Table 1: WTI, PPI, CPI Data Records of Pre-Pandemic, and Post-Pandemic Percentages

	Pre-pandemic (2015-2019)	Pandemic before vaccination (2019-2021)	Recovery after vaccination (2021-2022)	Forecast (2022 -2023)
WTI portfolio %	36.82%	35.64%	39.94%	28.00%
PPI portfolio %	29.54%	32.60%	32.04%	40.00%
CPI portfolio %	33.64%	31.76%	28.02%	32.00%
maximum return vs equal weighted	10.05% vs 9.22%	5.58% vs 5.31%	24.13% vs 21.25%	4.38% vs 4.15%
Saving	~1%	~0.3%	~3%	~0.3%

Figure 8 shows that because the optimal inventory portfolio diversifies non-systemic risks, it mainly reflects systemic risks, i.e., the inflation shock wave due to the COVID-19 pandemic. Percentage holding of WTI-correlated material shows that the shock wave first peaks in the energy market (in 2021-2022), while percentage holding of PPI-correlated material shows the ripple effect of the shock wave that starts to peak in the producer market in (2022-), with percentage holding of CPI-correlated material starting to pick up speed to reach peak.

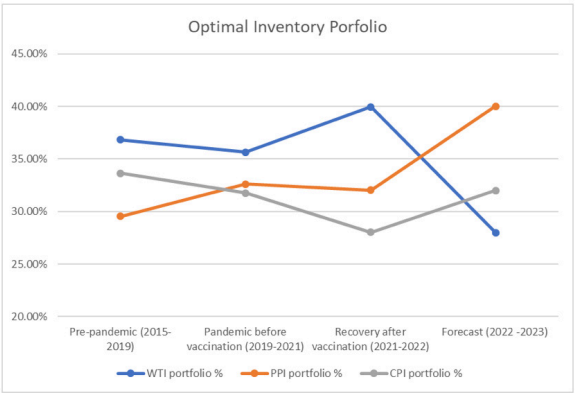


Figure 8: Optimal inventory portfolio links to market inflation shock wave due to COVID-19 (see Table 1)

In the above example, the market indices mainly reflect market systemic risks, with WTI the most volatile and with aggregate PPI and CPI the least volatile. Most individual businesses are connected to many market indices that reflect non-systemic risks of specific supply chain businesses and are volatile in between the two extremes. So, the expected saving of inventory portfolio optimization can be much larger than the results shown here. Future research can be to apply the method to a specific business, such that its non-systemic risks can be reduced and optimal balance among service, cash, and cost achieved.

Conclusion

Businesses are like ships floating in a vast ocean of the global economy subject to dynamic risk waves of supply and demand. Risk pooling is a statistical concept that suggests that supply chain variability is reduced if one can aggregate variability across locations, products, or time. Modern Portfolio Theory (MPT) offers a methodology to pool a diversified portfolio to minimize non-systemic risks to find a balance among services, cash, and cost of supply chain triangle. This paper formulates a balance of the supply chain triangle as a portfolio optimization problem, which maximizes savings of the inventory portfolio subject to constraints on inventory turns (service demand/inventory). This approach can further be used to achieve flexible operations by optimizing inventories for high-volume, low-variability materials to take advantage of economies of scale and/or optimizing inventories for low-volume, high-variability materials to pool the risks.

Acknowledgements

Thanks to Alex Tentopoulos and Vincent Lafrado from SUEZ WTS for their support in this science fair project.

References

1. Hubbard, R. G., & O'Brien, A. P. (2008). Economics. Upper Sad dle River, N.J: Pearson Prentice Hall.
2. How Climate Change Is Disrupting the Global Supply Chain https://e360.yale.edu/features/how-climate-change-is-disrupting-the-global-supply-chain
3. Coronavirus: How the pandemic has changed the world economy https://www.bbc.com/news/business-51706225
4. Covid-19 Related Inflation Surpasses 40 Year Record: How Long Will It Persist? https://www.forbes.com/sites/mikepatton/2022/04/13/covid-19-related-inflation-surpasses-40-year-record-how-long-will-it-persist/?sh=2da297eb1a82
5. TradingEconomics.com. https://tradingeconomics.com/
6. Simchi-Levi, D. (2013). Operations Rules: Delivering Customer Value through Flexible Operations. The MIT Press
7. Bodie, Z., Kane, A., & Marcus, A. J. (2011). Investments. New York: McGraw-Hill/Irwin
8. Markowitz, H. (1952). Portfolio Selection, The Journal of Finance, Vol. 7, No. 1, pp. 77-91
9. https://blog.arkieva.com/balancing-cash-cost-service-supply-chain-triangle/
10. PortfolioVisualizer.com. https://www.portfoliovisualizer.com/
11. WebDigitizer. https://apps.automeris.io/wpd/

Author

Charley L. Wan is an 11th-grade student at Central Bucks East High School. This science fair project was inspired by his

investment club, where he used portfolio optimization for the Wharton business school investment competition. He is passionate about integrating a growth mindset and neuroplasticity with business decision-making and organizational behavior.

DOI: 10.36838/v4i6.3

IJHSR
ijhighschoolresearch.org
■ RESEARCH ARTICLE

Constructing a Choice-Boundary/Growth Mindset Model Through EDA and EEG

Charley L. Wan
Central Bucks East High School, 2804 Holicong Rd, Doylestown, PA 18902, USA; charleywan2004@gmail.com

ABSTRACT: Decision-making is becoming increasingly essential to a person's success in life. There is a generational shift from passion to practicality. For example, while Boomers and Millennials focus on finding a job that interests them and cultivates their passion, iGen (the current generation) has their eyes on the money earned for the job.

This paper studied a student's different reactions towards learning and the completion of tasks in an educational environment and analyzed the relationships between energy consumption and the type of choice a student makes using Electrodermal activity (EDA) and Electroencephalography (EEG). The results yielded the conclusion that based on the amount of energy a student can provide for the task, their corresponding choice can be determined and the boundary in which that student works best can be uncovered.

To represent this relationship, a choice-boundary model is proposed where energy consumption is expressed as a Lyapunov function of choices. The model expresses possible directions a participant might take when introduced to a task and outlined possible pathways the participant might take in the process of completing the task. What this model and relationship imply is that in an educational environment, there is a happy medium between passion and practicality, in which a growth mindset can be cultivated, and where choices and their boundaries are well balanced while also allowing learning development.

KEYWORDS: Growth Mindset; Neuroplasticity; Choice; Boundary; Lyapunov function; Electrodermal activity (EDA); Electroencephalography (EEG).

■ Introduction

Decision-making is becoming increasingly essential to a person's success in life. In her book iGen, Jean Twenge distinguishes specific differences between teenagers during the Boomer, Millennial, and iGen generations. One of the main differences occurs in the way youths view life and its purpose. In an iGen'ers typical education, many students think of school more practically in terms of economic and social benefit. This practical view reveals that students today are making choices and setting boundaries in a drastically different manner when compared to their older counterparts. One example of this change is iGen'ers opinion on jobs. While Boomers and Millennials focus on finding a job that interests them and cultivates their passion, iGen'ers have their eyes on the money earned for the job.[1]

Schools have always advocated for an educational environment that cultivates growth mindset. A "growth mindset" thrives on challenge and sees failure not as evidence of unintelligence but as a heartening springboard for growth and to a challenge student's existing abilities. On the other hand, a "fixed mindset" assumes that character, intelligence, and creative ability are static givens which cannot change in any meaningful way; striving for success and avoiding failure at all costs becomes a way of maintaining the sense of being smart or skilled.[2] Growth mindset is supported by neuroplastic tendencies of the human brain. The human brain makes a variety of networks as a child continues through their lives from learning of new tasks and mastering skills.[3] The "softwiring" and "hardwaring" of the human brain are constantly changing. While learning of a new skill "burns" new knowledge into the brain's neural network, (softwiring) mastery of a skill reuses and consolidates the brain's existing networks.[4] (hardwiring)

Passion or practicality and growth mindset or fixed mindset are among many decisions a student has to make. In Stephen Covey's 8th Habit Book, he states:

"Between stimulus and response there is a space. In that space lies our freedom and power to choose our response. In those choices lie our growth and our happiness."

Each student has their own preferences on the best way to learn, so growth mindset does not have a defined structure. Based on the amount of energy a student can provide for a task, their corresponding choice can be determined and the boundary in which that student works best can be uncovered. By implementing the variables of choice and boundary, students' personal "growth mindset" zone can be derived from a structured model.[5]

The paper is outlined as follows. In the first part of the paper, neuroplasticity and growth mindset are measured by using Electrodermal activity (EDA) and Electroencephalography (EEG). EDA responses are tiny electrical changes on the skin which may indicate the body's response to stress. This data was collected by applying a low, undetectable, and constant voltage to the skin and then measuring how the skin conductance varies. This can be done with EDA devices that measure the electrical signal recorded by electrodes applied to the skin. EEG info: The Muse EEG headband contains seven electrodes, two on the forehead (prefrontal cortex), one behind each ear (temporal lobes), and three more electrodes as references. The brain waves recorded with this device are

DOI: 10.36838/v4i6.4

signal moving between 3 mental states of calm, neutral and active. [6]

While EDA is a measurement of growth in the body's autonomic nervous system, EEG is a measurement of growth in the body's central nervous system. In the second part of the paper, a choice-boundary model constructed, where energy consumption is expressed as a Lyapunov function of choices. The model articulates possible directions a participant might take when introduced to a task and outlines possible pathways that participant might take in the process of completing the task. What this model and relationship imply is that in an educational environment, there is a happy medium between passion and practicality in which a growth mindset can be cultivated, and where choices and their boundaries are well balanced while also allowing learning development.

Methods

EDA as a measurement of neuroplasticity/growth via Choice Substitution:

For the EDA experiment, the objective was to observe sensory substitution and the "growth" in the body's ability to adapt to increasing task difficulty. The experiment consisted of a simple call-to-action activity (Bop-It!) and tested a participant's ability to use senses other than hearing to complete the same mastery score (i.e., performing 100 actions with increasing difficulty without mistakes). Earmuffs were used to block out auditory sensing, while the Buzz Band was placed on the participants lips to "feel" the sounds, and a sound wave app was used for the participant to "see" the sound. With all the substitutions combined, each test recorded how long the participant was able to correctly perform the actions inputted by the different devices. An EDA fingertip sensor was used to record the micro voltage (potential energy) that was experienced by the stimuli. In summary, it took thirteen days: 20 minutes (10 minutes choice original, 10 minutes choice substitute) in each experiment to develop matching skills between the choice original and the choice substitute (see Table 1).

Table 1: EDA Measurements Between Original and Substituted Senses.

No	Choice Original (also indicator of game difficulty)	Choice Substitute
1	0.2779	0.3141
2	0.9002	0.8023
3	0.6666	0.6263
4	1.2811	1.3427
5	1.7437	0.777
6	2.0933	1.4067
7	1.5677	1.6456
8	1.5302	1.3096
9	1.9449	1.1925
10	1.4457	0.6973
11	2.0321	1.7335
12	1.652	1.357
13	3.068	2.8607

"EGG info": EEG responses are recorded brain waves on the surface of the scalp which may indicate the brains response to stimulation and cognitive energy consumed. This data was collected through the muse EEG headband.

EEG as a Measurement of Neuroplasticity/Growth via Choice of Exploration vs Exploitation :

In this experiment, three participants were tested for an exploration task and an exploitation task. In the exploitation task, the participants would play a video car driving video game in a fixed environment with fixed steering controls. In the exploration task, the participants would read a classical novel (i.e., Lord of the Rings). In the same day, both tasks were tested in intervals of 10, 15, and 30 minutes, with 5-minute rests in between. For six days, this process was repeated, to gather all the data (see Table 2).

Table 2: EEG Measurements Between Tasl Duration and Energy Amount

Index	min	read-rest	game-rest	(read-rest)-(game-rest)
1	10	3.575222	1.1382728	2.43694901
2	10	5.40662	3.841237	1.565383087
3	10	2.703986	2.9260308	-0.22204534
4	10	2.299625	0.6717958	1.627829454
5	10	4.33242	0.8908769	3.441542962
6	10	1.482907	1.5138055	-0.03089807
7	15	3.950033	2.076292	1.873740705
8	15	4.086221	3.4164465	0.66977425
9	15	1.809697	-1.2625531	3.07225026
10	15	2.939217	0.2917803	2.647436178
11	15	0.512136	-0.1227801	0.63491589
12	15	1.997477	0.6678265	1.32965047
13	30	3.891807	1.366901	2.52490638
14	30	4.059761	2.4450469	1.61471414
15	30	1.359963	2.7251717	-1.36520892
16	30	3.56596	3.3502243	0.215735363
17	30	1.116727	0.7195272	0.397199536
18	30	0.667903	1.1520632	-0.48416022

Results and Discussion

EDA as a Measurement of Neuroplasticity/Growth via Choice Substitution:

The choice substitution by interchanging inputs from the senses in the human body revealed a relationship between time and the magnitude of energy required to accomplish a task. Specifically, auditory inputs were substituted by visual and epidermal inputs (see Figure 1). The main objective, therefore, was to observe the familiarization and mastery of performing a task. As the experimentation continued, performance on both the original and substituted senses reached an approximately similar endpoint. When shown side-by-side, the progression of "mastery" in both tests displayed relatively identical movements in energy consumption (see Figure 2).

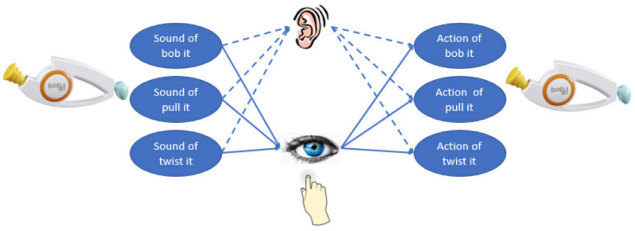


Figure 1: Diagram of EDA Experimentation. Senses with inputs of sound were substituted by sight and touch. The participant felt and viewed sound waves instead of listening to them. The changed inputs yielded outputs that highlighted the body's ability to adapt to difficult inputs and strengthen the new inputs to more efficient and accurate performance.

DOI: 10.36838/v4i6.4

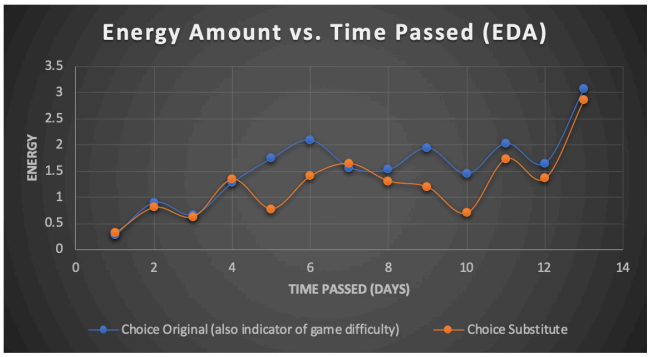


Figure 2: Relationship between Task Difficulty and Passage of Time in correspondence to energy usage. Both tests show an overall increase in energy consumption, and both tests converge to a relatively similar endpoint.

The underlying factor to the eventual conformity of both tests was choice. A participant could either perceive a task competitively or casually. The choice to perform well caused the intentional increase in energy usage in order to reach the required completion of the task (in this case, performing 100 actions with increasing difficulty without mistakes).

When exposed to a new task or assignment, the human body uses the multiple sources of inputs from the senses to gather background information, to observe, and to establish a foundation. For both tasks, the improvement follows a similar increase of energy consumption. Compared to the sensory inputs from the most familiar and efficient source(s) (i.e., hearing), the brain takes more time to adapt to slower and less efficient substituted sensory inputs (i.e., sight and touch). After familiarizing the procedures and mechanics of the alternative sources, the brain's performance on the task slowly approaches the energy level of the tests with no sense limitations.

EEG as a Measurement of Neuroplasticity/Growth via Choice of Exploration vs Exploitation:

In tandem with the EDA measurements, the measurements of EEG were also used to observe the changes in energy consumption for tasks over the passage of time. The objective of this experiment was to investigate the relationship between task duration and energy consumed. A key factor to highlight is the distinction between an exploration and exploitation task. An exploration task absorbs new information and converts it to familiar information over time, but an exploitation task takes familiar information and strengthens its performance over time (see Figure 3).

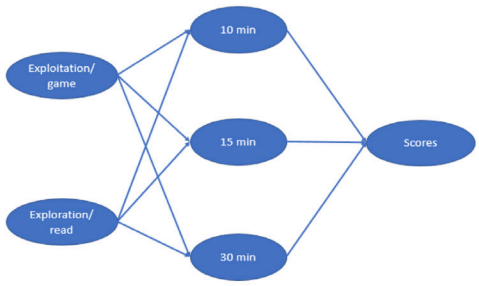


Figure 3: Diagram of EEG Experimentation. Exploitation task (game) and exploration task (read) are tested in 10 min, 15 min, and 30 min intervals with 5 min rests in between. The comparison yielded the conclusion that while each task starts at a different position, the amount of energy needed to complete both types of tasks eventually converges over time.

The difference between the two tasks is what a participant experiences at the beginning of the experiment. Since the exploitation task simply repeats a previously performed task, it will begin at a relatively low energy consumption and this consumption will stay in its low position. However, the exploration task constantly inputs new information; it will begin at a relatively high energy consumption. The nature of longer task duration displays an inverse correlation with energy consumption. The more familiar a participant is with the task; the less effort and exertion is needed to complete it. Although each task starts at a different position, the amount of energy needed to complete both types of tasks eventually converges (see Figures 4 and 5).

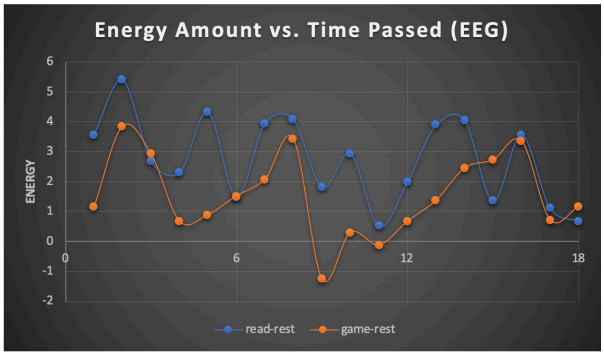


Figure 4: Relationship between Task Duration and Passage of Time in correspondence to Energy usage. Both tests have an overall decrease in energy consumption, and both tests converge to a relatively similar endpoint as well.

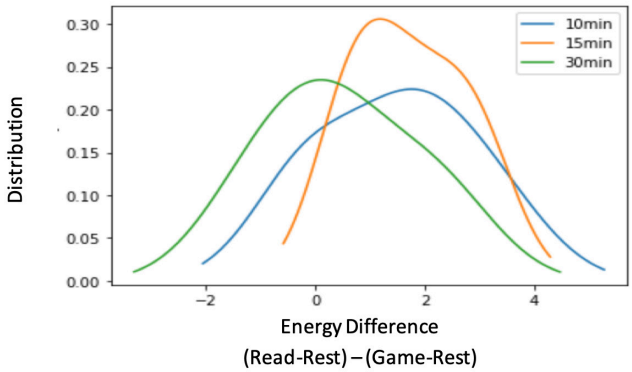


Figure 5: Distribution of Energy in Task Completion in correspondence to Task Duration. The graph shows that as task duration increases, the average energy usage decreases.

Choice and Boundary Model for Growth Mindset:

From the above experiments, it can be observed that for every choice made, a specific amount of energy is needed (i.e., time and effort). Yet, energy is limited by both its supply and its demand, which in turn set the boundaries for the choices that can be made. To represent the relationship in the above two experiments, a choice-boundary model is constructed, where energy consumption is expressed as a Lyapunov function of choices (see Figure 6). Specifically, the vertical axis is V, the energy consumed, the two horizontal axes are X1, task speed or the inverse of task duration, and X2 which is task difficulty. The Lyapunov function represents a model of a dynamic system. A stable dynamic system gradually loses energy (dissipation) and settles to the bottom of the bowl. An unstab-

le dynamic system gradually increases energy and particles will keep rising to the top of the bowl. Unlike a simple dynamic stable system which eventually reaches a stopping point, humans are complex systems; the life of humans can be imagined to be tan unstable dynamic system. The gradual increase in energy can be represented as, occasional pushes which keep the pendulum in motion. These occasional pushes are directly connected to choices.[7]

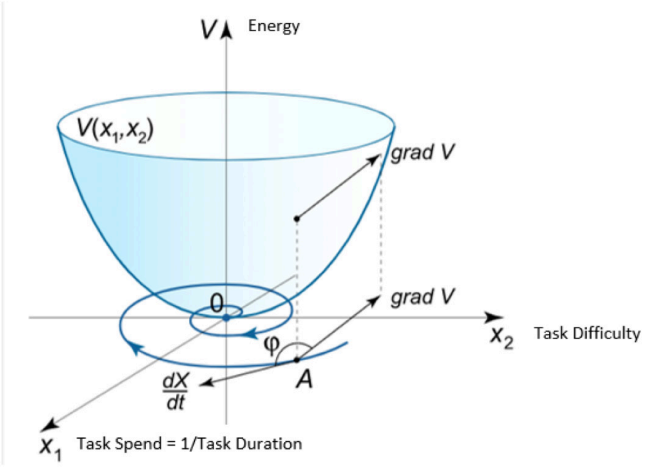


Figure 6: Diagram of Lyapunov Function, which shows that all stable systems eventually spiral down to a resting state.

The choice and boundary model expresses the direction of human behavior towards a growth mindset and mastery of tasks. One aspect of this model shows an upward and downward route. The upward route encourages the choice of increasing task difficulty which cultivates extended boundaries and additional opportunities. The downward route encourages the choice of task mastery via longer task duration with task completion more automatic and with less energy required.

This model acknowledges that there are specific characteristics in each choice that people could to be aware of. From the model, one who chooses to pursue one's passion (upward direction with law task duration and high task difficulty) is vulnerable to stress because the boundaries presented consume too much energy. Therefore, people may experience the feeling of lagging behind the required level of activity. A balance must be established between energy demanded and energy supplied in order to preserve a constant positive momentum in life. By keeping this momentum, one can prevent the issue of overworking.

On the other hand, with high task duration and low task difficulty one who chooses to focus on practicality (the downward direction) may be vulnerable to monotony because the boundaries presented consume little to no energy. Therefore, people may experience the feeling of disinterest or lack of exploration. In this case, the boundary limits a person's ability to discover new things among his or her small range of activity. A balance must be established between energy demanded and energy supplied to preserve a constant positive momentum, but also to prevent oneself from arriving at a dead end (i.e., reaching zero in the Lyapunov function).

This model can be applied in two different ways based off choices. The first way is to decide whether the task is constantly changing or if it is static. In relation to what was explained above, the tasks that change will most likely require a person to explore through trial-and-error and expose them to multiple high risk, high reward situations. Therefore, the best choice would be to adapt to the mindset of the upward direction. The increasing task difficulty, according to the model, demands more energy from the person, but allows him or her to find and pursue interests. If the task is already defined by a specific boundary, it will most likely require a person to perform more efficiently and expose them to routines. Therefore, the best choice would be to adapt to the mindset of the downward direction. The increasing task duration, according to the model, demands more stamina from the person, but allows him or her to become a master of specific skills.

The second application is choice blending. It is true that a person might not follow the model's recommendations to the extreme. There must be a happy medium in which both passion and practicality are implemented into daily life. A second representation of the model shows specific ranges where both choices can be blended. The ranges do not necessarily define a person's boundaries, but it does distinguish the places where the two dimensions of choice can be blended (see Figure 7).

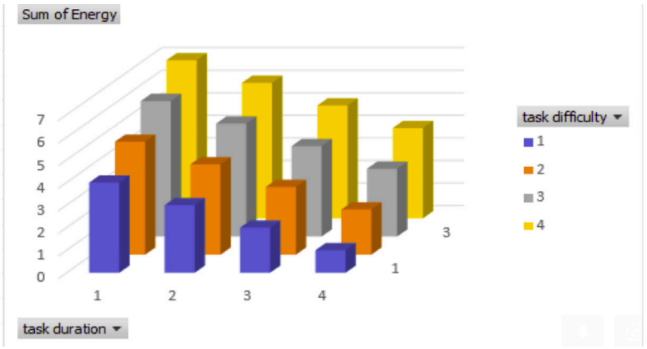


Figure 7: Choice blending of Choice-Boundary Model shows the different amounts of energy in terms of task duration and difficulty. The bar graphs also show that one energy requirement includes, but is not limited to four different ways of dispersing task duration and difficulty.

According to the bar graph, task difficulty changes by color (blue= low, orange= low medium, gray= high medium, yellow= high) and task duration changes from left to right in increasing order. With the two factors combined together, a total energy requirement can be determined. A task that is easy will require the same amount of energy in a short amount of time as is a task that is hard that takes a longer duration. For example, completing a task that requires an energy level of 4 presents four possible combinations of duration and difficulty. A student can decide which combination is best for him/her. Specifically, when presented with a task that demands high energy usage, a student may choose to increase the time of completion to reduce physical and mental stress while improving skill mastery. The opposite operates similarly. When presented with a task that demands low energy usage, a participant may choose to shorten the time of completion to reduce boredom while improving time utilization. Each student has his/her own abilities and perception of challenge; this model can be used to plan the strategy to implement a growth mindset.

DOI: 10.36838/v4i6.4

Conclusion

The ideal balance in which a student can develop a flourishing growth mindset is where energy is used and spent productively and wisely on the task. Given the limit of a student's effort, both the difficulty and duration of a task can be altered to provide enough interest for the student to be motivated intrinsically in the process of completing the task, but also enough flexibility to allow for the student to rest while continuing to encourage exploration on his/her own time.

The EDA data expressed the relationship in which the harder the task becomes, the higher the energy demand was for the participant to adapt to the challenges. The EEG data expressed the relationship in which the longer the task becomes, the lower the energy demand is. Thus, the participant can complete the task efficiently. Both cases showed the existence of growth in terms of learning of new skills and mastering of existing skills.

The model proposed can represent some of the many dispositions a participant may exhibit under a growth mindset. A task can be hard, but the participant can complete it in small steps. Similarly, a task can be long, but the participant can break the time into segments, each having its own specified difficulty. In both cases, the student is using the divide-and-conquer strategy, accumulating energy for increasing energy demand. On the other hand, a task can be easy, but the participant can become master of the skill and to improve energy utilization or complete it in a shorter amount of time to improve time utilization. A task can be short, but a participant can prolong task duration in order to explore more aspects of the task, or to extend its concepts to other disciplines. In this case, a student is using the short task to share duration and difficulty with other tasks.

In summary, the experiments and the model present possible directions a participant might take when introduced to a task and outline possible pathways the participant might take in the process of completing a task. What this model and relationship imply is that in an educational environment, there is a happy medium between passion and practicality, in which a growth mindset can be cultivated, where choices and their boundaries are well balanced while also allowing learning development.

References

1. Twenge, Jean M., (2017). IGen: Why Today's Super-connected Kids Are Growing Up Less Rebellious, More Tolerant, Less Hap py-- and Completely Unprepared for Adulthood (and What This Means for the Rest of Us). New York, NY: Atria Books.
2. Eagleman, D. (2017). The brain: the story of you. First Vintage Books edition. New York: Vintage Books.
3. Lee, Daeyeol. "Game Theory and Neural Basis of Social Decision Making." Nature.com, 26 Mar. 2008, https://doi.org/10.1038/nn2065. Accessed 15 Sept. 2021.
4. Dweck, C. S. (2006). Mindset: The new psychology of success. New York: Random House.
5. Covey, Stephen R. (2004). The 8th habit: from effectiveness to greatness. New York: Free Press,
6. Muse. (2018, June 25). A deep dive into brainwaves: Brain wave frequencies explained. Muse. https://choosemuse.com/blog/a-deep-dive-into-brainwaves-brainwave-frequencies-explained-2/
7. Frankl, Viktor E. (Viktor Emil), 1905-1997. Man's Search for Meaning; an Introduction to Logotherapy. Boston: Beacon Press, 1962.

Author

Charley Wan is an 11th grade student who attends Central Bucks East High School. Along with his scientific interests in both physics and chemistry, social sciences have piqued his curiosity in delving into the complexity of human brains and their interaction with the world. He aims to study computational neuroscience.

IJHSR
ijhighschoolresearch.org
■ RESEARCH ARTICLE

The Climate-Driven Adaptation of Mixotrophic Green *Noctiluca scintillans* in the Arabian Sea

Alisa M. D'Souza
Peekskill High School, 1072 Elm Street, Peekskill, New York, 10566, USA; alisamdsouza@gmail.com

ABSTRACT: Green *Noctiluca scintillans* (*Noctiluca*) is an unusually large mixotrophic dinoflagellate that forms widespread blooms in tropical coastal marine ecosystems. Its recent emergence as the dominant winter-monsoon bloom-forming organism in the Arabian Sea has been attributed to the unique ability of its endosymbionts, *Protoeuglena noctilucae*, present within its symbiosome to photosynthesize more efficiently under suboxic conditions. What also distinguishes green *Noctiluca* from most mixotrophs and other bloom-forming organisms is its ability to persist and thrive as large blooms, even under the most unfavorable nutrient conditions. In this study, we report the results of 1) nitrogenous nutrient enrichment experiments undertaken with a laboratory strain of green *Noctiluca* isolated from the Arabian Sea, and 2) nitrogenous nutrient uptake experiments undertaken with natural populations of green *Noctiluca* sampled off the coast of Oman. Both investigations revealed that this organism has a greater preference for regenerated nitrogenous nutrients, i.e., urea and ammonium (NH_4), as compared to new nitrogen, nitrate (NO_3). What was particularly surprising about this mixotroph, however, was its ability to grow and survive for extended periods, even in the absence of nitrogenous nutrients. In the field, green *Noctiluca* bloom outbreaks were invariably preceded by extremely high concentrations of urea and NH_4. We also observed that irrespective of the nitrogenous nutrients available to them, green *Noctiluca* cells always accumulated large amounts of NH_4 within their symbiosome, which resulted in sharp increases in seawater NH_4 concentrations (30 to 200 μM) following their demise.

KEYWORDS: Earth and Environmental Sciences; Environmental Effects on Ecosystems; Climate Change; Mixotrophs; *Noctiluca scintillans*.

■ Introduction

The Arabian Sea (AS) is one of the world's fastest-warming ocean ecosystems. It harbors a permanent oxygen minimum zone (OMZ) that is roughly three times the size of Texas.[1] Warming and circulation processes in the AS are under the control of monsoonal winds which change their direction from southwesterly in summer (June to Sept.) to northeasterly in winter (Nov. to Feb.). The strength of these monsoonal winds is controlled by land-sea pressure gradients, hence the warming and cooling of the land and sea. Therefore, circulation processes and biological productivity in the AS are especially susceptible to global warming and climate change.[2,3] The AS marine and the surrounding terrestrial land biomes are habitats for a rich and diverse array of organisms, from phytoplankton and zooplankton at the base of the food chain to fish and larger marine mammals; it sustains the livelihoods of millions of humans living along the coasts of bordering nations of Oman, Iran, Pakistan, and India (see Figure 1 for AS region map). As such, the AS provides an important source of food and economic opportunity for the littoral populations (both tourism and fisheries are a substantial part of coastal livelihood). Recent research is providing conclusive evidence of alterations in the dynamics of this traditional ecosystem of the AS that impact not only the oceanic processes and organisms that live within the sea but the base (and consequently the rest) of the food chain, with serious socioeconomic implications.

Figure 1: The Arabian Sea region on the map.

Altered nutrient levels and ratios, processes involving certain nutrients and lower trophic level organisms, and food chains are among some of the early ramifications of climate change and other anthropogenic-related pressures in oceans around the world.[4] What makes the AS particularly susceptible to global warming is that it is inherently warm and nutrient impoverished, and its biological productivity is tightly coupled to the strength of monsoonal winds and circulation processes. The OMZ expansion can be largely accredited to 1) ocean stratification, as it inhibits the diffusion of atmospheric oxygen (O_2) into the water column and 2) the overall warming of the sea, because as waters warm, the solubility of O_2 decreases, resulting in a reduced dissolved O_2 content.[5] As a result of O_2 loss, bacteria that rely on dissolved O_2 in seawater rely on the O_2 from the NO_3 dissolved in seawater,

DOI: 10.36838/v4i6.5

causing subsequent NO_3 loss and NH_4 accumulation as the conventional nitrogen cycle is disrupted. In the AS, the accumulation of NH_4 in seawater is compounded by inputs from rivers. These nutrient inputs include both NH_3 and urea from fertilizer runoff, wastewater or sewage treatment plant effluents from large coastal cities, such as Karachi, Mumbai, Aden, Salalah, and Muscat.[6]

Typically, winter monsoonal winds coming from the Himalayan mountains cool the surface of the AS and the surrounding landmasses, which promotes circulation through intensified cooling of surface waters and convective mixing. However, with increased Earth temperatures, the winds coming off the Himalayan mountains have become warmer because of the decline in snow accumulation in the Himalayan mountains. These warmer winds dampen convective mixing, warm the surface waters and further stratify the upper water column, causing a decline in nutrients brought up to the surface waters.[3] Additionally, warmer waters are already observed as climate change increases land surface temperatures, and this, as mentioned before, reduces the amount of O_2 the AS can hold and expands the OMZ.

In addition to the abnormal patterns of convective mixing, water column stratification and OMZ expansion, the increasing presence of harmful algal blooms (HABs) forms yet another threat to the sea and the surrounding area. HABs are caused by the rapid growth of algae composed of either cyanobacteria, dinoflagellates, or diatoms that coat water surfaces; their impacts can range from being small disturbances of the food chain to massive fish mortality when they are toxic. In the AS, the spatial distribution and growth of HABs are influenced by climate change, freshwater sources, monsoonal patterns, changes in convective mixing patterns/currents, seasonal temperatures, wind direction, and velocity.[6-9]

HAB transportation and growth are also heavily determined by the nutrient levels and temperatures of the water, tying in climate change and human activities. Natural dust deposition, monsoons, and cold eddies can impact nutrient levels as well. Climate-related changes can also break down natural biogeographical barriers with intensified storms, a higher frequency of hurricanes, and more severe floods. These barrier breakdowns can lead to increased HAB outbreaks and transportation as well.[10] HABs have been seen across the Arabian Sea since 1908, and cause several problems for power plants, local schools and children, the region's tourist rates and economy, and the environment and ecosystem that host the blooms.[6]

HABs can affect the surrounding area in 3 major ways: the water quality, economy/socioeconomics, and of course the ecosystem. Water quality: The water quality of an area can be compromised as the blooms thrive and potentially clog desalination and other water treatment plants. The toxins that some HABs release can harm, or even kill, fish and other animals, including humans if they consume something that has been contaminated. Not all blooms are toxin-producing, though, and can instead harm water quality and the ecosystem by blocking light from passing through the surface, clogging fish gills and desalination plants, and providing burdens for the traditional organisms in the ecosystem through its dominance[11]

Economy: This can include tourism and fisheries. For tourism, HABs are unattractive with their foul smell, and their coating of water surfaces can prevent recreational swimming, fishing, and other activities. For fisheries, HABs are linked to fish mortalities and less fish diversity (either through deoxygenation or toxin releases), various health issues in humans, as well as an increase in fish prices because of less availability.[1] The presence of HAB– and their impacts- thus provide challenges for the area's socioeconomic system, especially for those who rely on tourism and fishing for livelihood, which are primarily the countries and coastal cities that border the AS (known as "rim countries").[9]

Ecosystem: If HABs persist, they have the potential of "taking over" the region's ecosystem and worsening their impacts. With an unusual abundance of phytoplankton at the base of the food chain comes disruption of the balance of the rest of the ecosystem; the amplification of toxins that some HABs contain (as previously mentioned) can also harm the ecosystem's health. HAB decay draws out oxygen from the water, and HABs thrive off certain nutrients, so there is also a nutrient disruption that affects other organisms and processes in the ecosystem.[12]

One of the most prominent HABs in the Arabian Sea is green *Noctiluca scintillans* (*Noctiluca*), which has had an increasing presence there during the winter monsoons since the early 2000s (see Figures 2a-c). There is also the red *Noctiluca scintillans* (red *Noctiluca*), which is a heterotrophic grazing organism, that consumes microzooplankton and does not photosynthesize. Red *Noctiluca* thrives in higher salinity waters, at 10-25°C.[13]

As a mixotroph, green *Noctiluca*, in contrast to the red variety, has a competitive advantage over most other lower trophic organisms, because it can sustain itself 1) by feeding on other organisms because of its innate heterotrophic trait, and 2) through photosynthesis and nutrient cycling provided by its autotrophic endosymbiont, *Protoeuglena noctilucae*.[14]

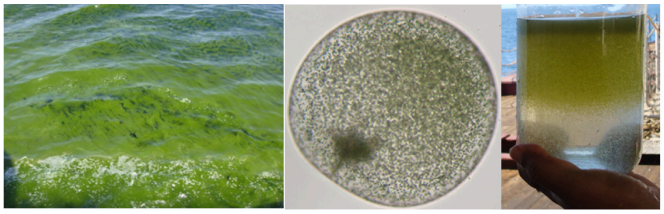


Figure 2a: Bloom image of natural populations of *Noctiluca.*
Figure 2b: A *Noctiluca* cell under the microscope with algal endosymbiont *P. noctilucae.*
Figure 2c: Sampled *Noctiluca* from a bloom.

It is interesting to note that *P. noctilucae* evolved in a low oxygenated (oxygen, O_2), highly carbonated (carbon dioxide, CO_2), environment 1.3 billion years ago, whereas *Noctiluca* is a rather new organism (dating back to only a few hundred years ago). *Noctiluca* thus depends on this *P. noctilucae* to help it adapt to a changing world.[1] These two organisms live in a mutualistic symbiosis, and the unique nutrition acquisition and recycling of *P. noctilucae* allows *Noctiluca* to survive and

thrive in conditions where other organisms would not. The mixotrophy of *Noctiluca* has been studied under variable prey types and light concentrations, showing that *Noctiluca* can sustain life for prolonged periods with its endosymbionts. The study showed that without an external prey, *Noctiluca* growth rates were less than when fed with a prey.[14] *Noctiluca* endosymbionts also had carbon fixation rates between 25 and 300% higher in waters with low dissolved O_2 concentrations (i.e., near hypoxic conditions) than in ambient O_2 concentrated water. These results point to *Noctiluca* having a competitive advantage over other (non-mixotrophic) marine organisms which cannot keep up with the trend toward low O_2 levels in the AS.

However, there is no clear consensus regarding the specific environmental drivers for the annual blooms of green *Noctiluca* in the AS. For instance, a study conducted in the Northeastern AS (NEAS) found that the blooms of *Noctiluca* are not related to hypoxia, but are rather stimulated by intensified surface water stratification due to warmer waters from climate change.[15] The authors used in situ and satellite algorithm data to analyze bloom locations and cell abundance populations of *Noctiluca* and diatoms. In their analysis of water column environmental characteristics (temperature, salinity, dissolved O_2, *chl-a*), the authors found neither hypoxic nor anoxic waters in the sea, but rather comparatively high dissolved O_2 levels. This refuted the link between *Noctiluca* and low O_2 levels, even saying that "the suggestion by Gomes *et al*. (2014) that this linkage [between cultural eutrophication and blooms of *Noctiluca*] also can extend far out into oceanic waters is questionable". It may be noted that, unlike the contrasting studies, these findings were not based on *Noctiluca* with photo-physiological experiments but casual presence or absence of *Noctiluca* at different locations measured from space and few shipboard measurements. *Noctiluca* are free-floating organisms and can be easily transported by ocean currents to a different location from their point of origin.

Another study conducted in the NEAS also refuted the possibility of *Noctiluca* being tied to hypoxia, as well as sewage/anthropogenic outputs as a foundation for the blooms. The authors stated that convective mixing and stratification arising from climate change is the bigger and more likely issue than hypoxia, and hypoxia was found to be nonexistent in the sea.[16] Contradicting methods and results of studies on *Noctiluca* blooms prevent a full and ubiquitous understanding of overgrowth, specifically as they relate to causes and effects. These research gaps are not exclusive to *Noctiluca*; research has found that there are gaps in current knowledge about the overall effects of climate change and climate change-related pressures on marine ecosystems, most notably in responses to ocean acidification and O_2 levels.[17]

Despite the dispute over the origins of *Noctiluca*'s domination, it is clear that its relatively recent abundance during the winter monsoons has led to a shift in the base of the food chain (traditionally dominated by diatoms) and further depletion of O_2 in the sea, through its rapid decay and decomposition by bacteria.[1] These studies also show that as the sea may be shifting towards scenarios that are often not to the aid of conventional organisms, whatever they may be, green *Noctiluca* instead proliferates as a mixotroph.

To help better understand the environmental drivers behind the massive blooms of *Noctiluca*, a study was conducted from the end of August 2019 to February 2020, using previously collected mentor-provided research data from Oman and student-done lab data and analysis at the Lamont-Doherty Earth Observatory. Its research questions were: what are the nitrogenous nutrient preferences that allow *Noctiluca* to survive for extended periods? Does mixotrophy offer *Noctiluca* a competitive advantage for growth and survival over other phytoplankton especially when nitrogenous nutrients are unavailable? The goal was a better understanding of the blooms and future consequences for the ecosystem. This would help solve the problem of dangerous HAB overgrowth disrupting marine food chains, threatening fisheries (which increases economic pressures and food insecurity), and harming the water quality of oceanic ecosystems.

More specifically, this study aimed to examine how green *Noctiluca* is impacted by changes in nitrogen types, and which nitrogen type was preferred (new versus recycled/regenerated), because nitrogen pertains largely to other algae bloom growth and the changing of ecosystems. Embedded in the broad purpose of this study was to draw attention to the role of cultural eutrophication and whether NH_4 and urea from agricultural fertilizer runoff and sewage treatment plans could be contributing to *Noctiluca* bloom outbreaks. Concurrent with typical HAB behavior and research on other HABs, I hypothesized that *Noctiluca* grows faster in the presence of NH_4 and urea than in naturally occurring $NaNO_3$. I also hypothesized that in addition to the rapid growth of *Noctiluca*, its propensity to accumulate NH_4 made its presence as blooms unfavorable for the growth of other phytoplankton.

Statement of Purpose/ Societal Relevance:

To further note the importance of the results of this study and the impact *Noctiluca* has, especially in a changing environment, the focus of this section will be the societal relevance and connections of these laboratory and field results to the world *Noctiluca* and HAB research. Firstly, one must take worldwide HABs into consideration, both including and asides from *Noctiluca*. *Noctiluca* predominates mainly in warm, tropical waters. Blooms of *Noctiluca* appear with an increasing presence in the Yellow Sea, East China Sea, and the Arabian Sea/Oman region, and interestingly enough save proven to be detrimental to ecosystem health in all three of these regions.

The East China Sea *Noctiluca* HABs, though not inherently toxic, release large amounts of NO_4 as they consume toxic algae. High levels of NO_4 can be toxic to other phytoplankton. These blooms of *Noctiluca* can affect the waters as a traditional HAB does, and the reduced O_2 levels from aerobic respiration further threaten marine life. Although it is not certain, this phenomenon is believed to be linked to anthropogenic activities and outputs, as nutrients from sewage runoff fuel blooms further, similar to the growth patterns of *Noctiluca* in the AS.[18] Similarly, the Yellow Sea faces blooms of *Enteromorpha prolifera*, where the algae are also not toxic, but the sheer size and development of the blooms over the sea is detrimental to eco-

DOI: 10.36838/v4i6.5

system health. This bloom is yet another example of sewage outputs and nutrient disruptions leading to negative impacts on a body of water and its surrounding area due to subsequent HAB ououtbreaks[9] Perhaps more widely known are the blooms of Florida and Mexico, Massachusetts, and the Long Island Sound, composed of cyanobacteria and toxic dinoflagellates. The blooms around the Florida/Gulf of Mexico region stem from anthropogenic septic tanks and fertilizer runoff; those in Massachusetts have sources such as warmer waters and disrupted nutrient levels. Similarly, the suffocating Long Island Sound "red tide" bloom is linked to warmer waters and increased nutrient concentrations, such as nitrogen, phosphorus, and silicate.[20]

As discussed above, major HABs are linked to the same causes, what do all of these bloom examples mean for the world, and why do they matter? These HABs, much like green *Noctiluca* in the AS, often draw out the O_2 levels in the water and, because they live off of output nutrients, disrupt the nutrient levels further. In general, blooms are linked to low O_2 in the region, leading to the suffocation of marine animals and the detriment of other organisms such as lobsters, crabs, squids, and coccolithophores, especially paired with ocean acidification.[21] As aquatic species are put at risk, the humans in the surrounding areas suffer a loss in food sources and economic opportunity, especially in areas that heavily rely on the sea for tourism, fishing, or similar activities. HAB-contaminated water can also lead to various illnesses and poisoning from toxins, limiting the amount of non-polluted water available, as rivers, lakes, streams, and ponds are already filled with chemical outputs.

These are only a few examples of algae blooms; HABS coat the surface layers of water globally, and while they have been noticed throughout history (dating to 1908 as previously stated), knowledge is even more imperative as existing nutrient ratios and climate conditions worsen. The described ecosystems suffer calamity from the blooms, and all of them are linked to some form of pollution/anthropogenic outputs and warmer waters. Knowledge on the formation and presence of some HAB blooms, the effects they have on us humans, and information on how to limit the recurrences of the blooms should be well spread to create a more informed and conscientious society. This study was done to help increase research on the now annual blooms of *Noctiluca*, as contradicting research, as well as the ubiquity of *Noctiluca* and other HABs themselves, prevent consensus and provide concern for the future if not remedied.

Methods

Study Area:

Student laboratory work ("we"; "I") was done in the Lamont-Doherty Earth Observatory research unit of Columbia University, where two labs were used throughout the experiment. One was for most of the experiment (incubating cells, counting cells, FIRe, data, and graph analysis), and the other was for more of the "prep" work (such as filtering and microwaving seawater). My mentor collected data from sampling in the field (off the coast of Oman) during the Jan-Feb *Noctiluca* bloom of 2018 which allowed me to compare my laboratory results with the field observations. These data included microscopic counts of *Noctiluca* and other phytoplankton, nutrients (NH_4, urea, $NaNO_3$, phosphate), and chlorophyll.

The area sampled by researchers during the time of collection (23rd Jan-10th Feb 2018) is known for having stable waters, abundant nutrients, optimal growth temperatures for *Noctiluca* within the 5 months of *Noctiluca* blooms (roughly 26-27 ºC), and good salinity because of the antecedent winter monsoonal period. Samples were gathered from the peak *Noctiluca* bloom, and Figure 3 shows the details of the collection. 3A shows the chlorophyll levels, and the red areas indicate a high level of chlorophyll, which determined where the researchers sampled. 3B shows cyclonic eddies in the red swirls. As they move in a counterclockwise motion, this is a cyclonic eddy that brings up nitrogen-rich and low-oxygen waters, which prompted researchers to examine how the nutrients impact *Noctiluca* growth, which was then further proved by the lab results.

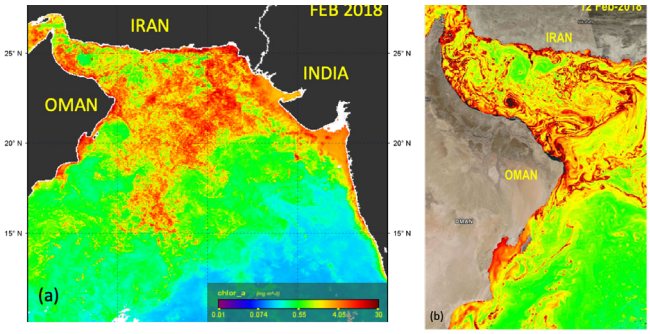


Figure 3: *Noctiluca* blooms as viewed from space by the Ocean Color Satellites (a) Monthly composite- NASA MODIS-Aqua (ocean color satellite) and (b) Daily image- NOAA VIIRS.

Cell Culture and Counts:

As an intern, I used lab-cultivated cells of *Noctiluca* previously collected from the Arabian Sea and placed them into 4 media types, 3 of which were a form of nitrogen and 1 was a control group. We used NH_4 chloride, sodium nitrate ($NaNO_3$), and urea as the tested variables, with urea and NH_4 as "regenerated" nitrogen types and $NaNO_3$ as "new" nitrogen; microwaved filtered seawater was the control to compare "normal" *Noctiluca* growth versus the nitrogen-impacted cells. CAUTION: Low concentrations of each of the hazardous chemicals (the tested/independent variables) were used, so there was no safety risk. Although, in the case of overexposure, irritation can occur to the eyes, skin, and respiratory system. NH_4 chloride can be the cause of all three, $NaNO_3$ for the eyes, and urea for the respiratory system. Chemicals were properly disposed of using mentor and SDS guidance, and nitrile gloves and standard laboratory attire (long pants, closed shoes, hair pulled back) were worn. Each nitrogen type (and the control) was used as a media in which the cells were allowed to grow for the duration of the experiment. The experimental flasks were labeled based on their nitrogen content, with three replicate treatments (3 separate flasks per nitrogen type) each labeled from A-C (Figure 4). The flasks were not hooked up to any gas chambers, as the only variable in the experiment was the nitrogen type that was in the media. In total, twelve bottles were filled with their respective

media and placed in an incubator maintained at around 27 degrees Celsius with a 300 mE m^{-2} s^{-1}, 14:10 (Light: Dark) cycle. When ready for analysis, we took cell counts and photosynthetic electron transport rates to assess the cell health and growth over four sampling days, and below are the details of the measurements.

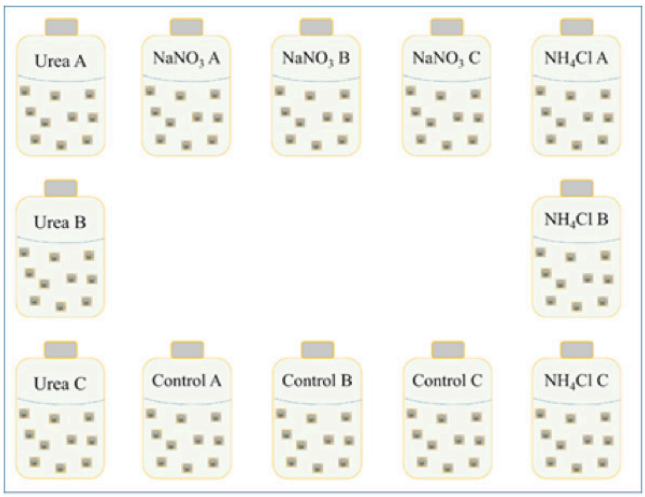


Figure 4: Experimental bottle setup of the cultures of *Noctiluca*; NH_4Cl represents NH_4 chloride, $NaNO_3$ is sodium nitrate, and urea is urea.

For the lab experiments: *Noctiluca* was enumerated manually in a 5mL pipette with the naked eye on sampling days.

Field: *Noctiluca* cells were counted after being collected from the bloom.

Changes in Noctiluca and Protoeuglena photo-physiology using Fluorescence Induction and Relaxation (FIRe) measurements:

At the start and throughout the experiment, I measured the photo-physiology and growth rates of *Noctiluca* and the free-living endosymbionts. For *Noctiluca*, 5 cells were picked up carefully with a dropper, washed with filtered seawater, and then transferred into 3ml of media. In the case of *P. noctilucae*, we used 3ml of only media from which all the *Noctiluca* cells were carefully removed. The FIRe provides a measure of bulk fluorescence (F_m) which can be related to the chlorophyll content of the *Noctiluca* cells or the endosymbionts that are released into the medium. Variable fluorescence (F_v/F_m) provides a measure of the health of the cells.

After the samples were processed in the FIRe, the cells of *Noctiluca* and that of the *P. noctilucae* were filtered on Whatman brand glass filters separately. After filtration, the filters were carefully stored in labeled histocaps in a -80°C freezer until the time of analysis. When ready for analysis, the filters were placed in a small 5 ml tube that contained cold 90% acetone to extract the chlorophyll from the filters. CAUTION: acetone can irritate the eyes, skin dryness or crack in the case of repeated exposure, drowsiness and dizziness from vapor exposure, and– in high concentrations– harm to the nervous system. Students took precautions by doing all the transfers of tilters into acetone in a fume hood and did not experience any of the symptoms of overexposure or irritation related to the acetone.

Results

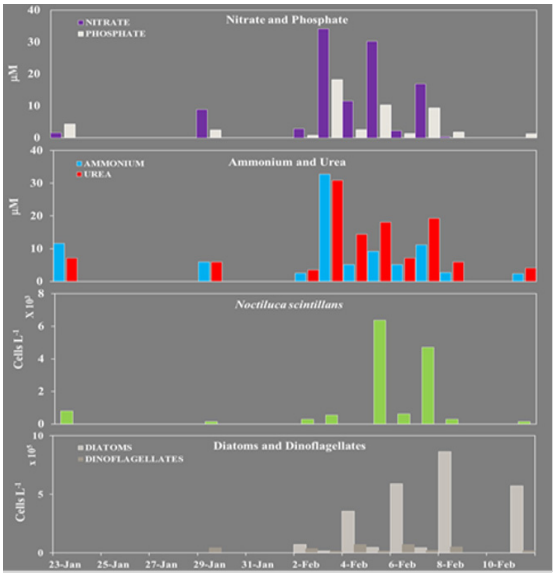


Figure 5: *Noctiluca* bloom peaks with increases of NH_4 and urea in the water column, and there was an increased diatom presence after *Noctiluca*'s demise.

Figure 5 depicts the cell counts from the field studies in Oman of populations of *Noctiluca* and other diatoms and dinoflagellates, against the $NaNO_3$, phosphate, NH_4, and urea (in µm) of the area surrounding the blooms. The collection was from January 23 to February 10, 2018.

The field ecological studies undertaken along the coast of Oman revealed peaks of *Noctiluca* blooms (of 4 and 6-7 x 10^3 cells per liter) when there was an increase in NH_4 and urea in the water column (around 30µm). We also see that after the demise of *Noctiluca* blooms, there is an increased presence of diatoms. Diatom presence was around 6 cells per liter (February 6) after the demise of the 6 cells per liter *Noctiluca* peak, about 8 x 10^3 cells per liter (February 8) after the demise of the 4 x 10^3 cells per liter *Noctiluca* peak, and about 5 x 10^3 cells per liter just around when *Noctiluca* existed at near 0 cells per liter levels (February 10). Other dinoflagellates (represented with the darker grey) do not seem to be particularly affected by the presence of *Noctiluca* or NH_4 and urea.

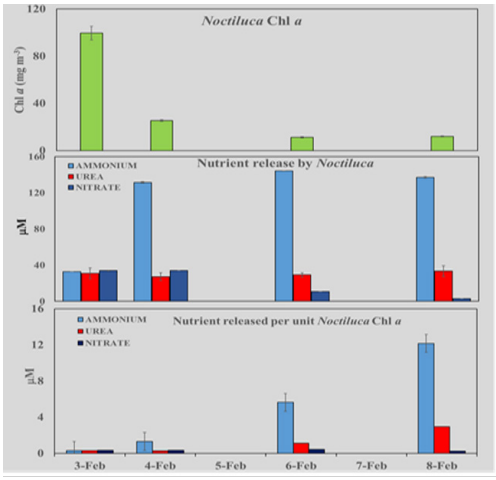


Figure 6: *Noctiluca* cells released a substantial amount of NH_4 upon their demise.

DOI: 10.36838/v4i6.5

The chlorophyll-a (*chl-a*) content of *Noctiluca* about the nutrient concentrations of NH_4, $NaNO_3$, and urea by *Noctiluca* (measured in μm), as well as the nutrient release in comparison to the unit chl levels of *Noctiluca*.

Figure 6 of the *Noctiluca* cells collected from the field bloom in Oman shows that the cells released a substantial amount of NH_4 (that accumulates in their central cytoplasm) upon their demise and decay. Before *Noctiluca*'s decay, released NH_4 was as nearly as high as 160μm, whereas released urea and $NaNO_3$ were only at most near 40μm.

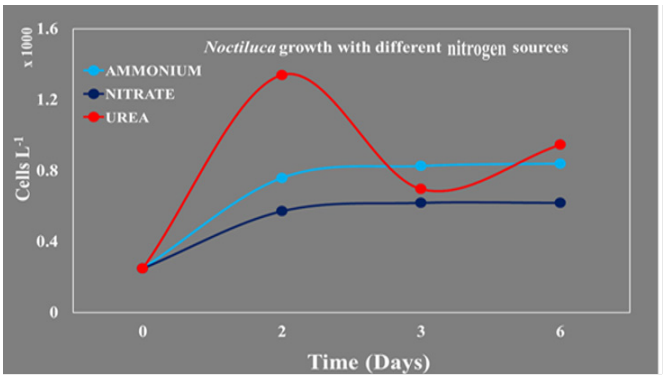


Figure 7: The laboratory cell counts of *Noctiluca*, under each of the three, tested nitrogen variables (NH_4, $NaNO_3$, and urea) in cells per liter, over time (each of the sampling days).

In the laboratory culture experiments, cells treated with urea generally had the highest growth, while $NaNO_3$ witnessed the least. Urea-treated cells show a maximum growth of over 1.2 x 10^3 cells per day, whereas NH_4-treated cells and urea-treated cells only reached a maximum of about 0.8 and 0.6 x 10^3 cells per day, respectively. Though not included in the graph, the control group faltered on the last day (Figure 7).

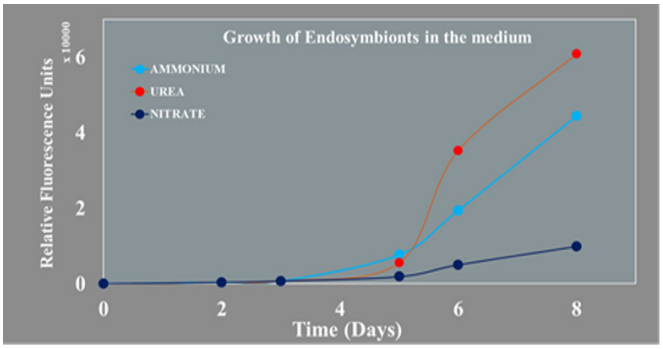


Figure 8: The growth of the endosymbionts, *Protoeuglena noctilucae*, is measured by relative fluorescence, in each media type (the variable nitrogen types: NH_4, $NaNO_3$, and urea) over time (each of the sampling days).

Figure 8 shows that when grown in urea and NH_3, the endosymbionts left the host *Noctiluca* cells and grew at a significantly better rate than when grown in $NaNO_3$. On day 8, the relative fluorescence of *Noctiluca* reached 6 units when treated with urea, 4 units when treated with NH_4, and not even 2 units when treated with $NaNO_3$. When paired with Figure 8, we see that both the host and the endosymbionts have higher growth rates in the presence of urea and NH_4.

Discussion and Conclusion

These experiments showed that 1) *Noctiluca* has a preference for "regenerated nitrogen" (urea and NH_4) in comparison to "new nitrogen" (nitrate) both in the lab and in the field, and 2), *Noctiluca* was able to grow and survive for extended periods, 3) green *Noctiluca* cells accumulated large amounts of NH_4 within their symbiosome, which resulted in sharp increases in seawater NH_4 concentrations (30 to 200 μM) following their demise. Our findings suggest that *Noctiluca* has a stronger preference for regenerated nitrogenous nutrients, i.e., urea and NH_4 over $NaNO_3$, in terms of its growth and NH_4 accumulation.

Figure 5 suggests that the *Noctiluca* blooms are triggered by NH_4 and urea and that diatoms and *Noctiluca* may live in an inverse relationship, where *Noctiluca* may be able to flourish during times of low diatom presence, and vice-versa is true for diatoms. When you pair Figure 6 with Figure 5, which showed that *Noctiluca* blooms are succeeded by diatoms, we can connect this to the stability of the water column and its relation to the nutrient outputs. If the water column is not stable (meaning there is more convective mixing), that would make the nutrients less concentrated in the area and therefore allow for the diatoms to flourish after *Noctiluca*'s demise.

As for the laboratory results, Figure 7 depicts *Noctiluca* growth being most efficient in the urea and NH_4 treatments, coincident with the growth responses of the endosymbiont *P. noctilucae* in Figure 8. What is important to note about Figure 8 is that those endosymbionts measured were from the media, indicating that the endosymbionts may prefer the external conditions to what is provided inside of the host. Particularly surprising about this mixotroph and the results as a whole is the ability of *Noctiluca* to grow and survive for extended periods even in the absence of nitrogenous nutrients. It is possible that *Noctiluca* rejected the endosymbionts, though, considering the advantages *P. noctilucae* provides for the host *Noctiluca*, it may not be so.

These results also lend credence to the idea that urea and NH_4 released from land-based activities by humans may be contributing to outbreaks of *Noctiluca*. Further studies will be needed to ascertain this hypothesis.

As mentioned previously, contradicting research is a common issue in marine and HAB research, so steps should be taken to limit the conflicting results from different studies, both through focused laboratory experiments on *Noctiluca* and other HABs. Having a plethora of "significant results" from multiple studies that lead science in miscellaneous directions is no better than only a few elements of research. The short time frame of this laboratory work may provide a limitation specifically for this research study. Moving forward, we would like to reassess the patterns noticed by including more sampling dates to assess patterns over a longer period—potentially 2-3 months, like that around the duration of natural *Noctiluca* blooms—as well as evaluate the different methods used in HAB and *Noctiluca* research and see if implementing different methodologies to this study would lead to a significantly different outcome. Additionally, the field and lab experiments' observations and data provided valuable insight

into *Noctiluca* and its endosymbionts, and how growth and physiological processes are affected by changing oceanic conditions and nitrogenous levels, but stable isotopes should also be taken into consideration to examine more *Noctiluca*'s internal processes.

HABs coat the surface layers of water globally, and while their overgrowth has been observed throughout history, knowledge of their impacts on ecosystems is even more imperative as existing nutrient ratios and climate conditions worsen. The described ecosystems suffer calamity from the blooms, and all of them are linked to some form of pollution/anthropogenic outputs and warmer waters. As for *Noctiluca*, the sources of the nutrients studied, including urea and NH_4, could include anthropogenic activity, such as fertilizer runoff and effluents from wastewater and sewage treatment plants.

The year-on-year increase in *Noctiluca* blooms in the AS appears to also be tied to climate change, because of the relationship between the OMZ and water warming, then the OMZ and the nitrogen fluxes, and then the nitrogen changes within *Noctiluca*'s responses. Knowledge on the formation and presence of some HAB blooms and the effects they have on us is well-studied and information should be well spread to limit the recurrences of the blooms. This is especially important for the populations around the AS, who need to be informed of the impact of their daily lives on the environment, and consequently themselves with the occurrence of *Noctiluca*. Together the scientific community, the public, and governmental agencies can develop plans to prevent further damage to the delicate balance of the Arabian Sea and local land ecosystems.

Acknowledgments

I would like to thank Lamont Doherty Earth Observatory at Columbia University for the opportunity to conduct this research. I am very grateful to Ms. Kali McKee, Dr. Helga does Rosario Gomes and Dr. Joaquim Goes for mentoring me during this study, and they were all a great help in both this experiment and previous ones. Additional thanks to Mrs. Erum Hadi, Science Research Teacher at Peekskill High School, for guiding me in my work throughout the school year. Thank you also to Khalid Al-Hashmi, Sultan Qaboos University, and the Ministry of Agriculture and Fisheries Wealth, Sultanate of Oman for their help with this work and field data. We are grateful to NASA Applied Sciences, the Gordon Betty Moore Foundation, and the Sultan Qaboos Cultural Centre for their support of this research study.

References

1. do Rosário Gomes, H.; Goes, J.; Matondkar, S.G.P.; Buskey, E.J.; Basu, S.; Parab, S.; Thoppil, P.; Massive outbreaks of Noctiluca scintillans Bloom in the Arabian Sea due to the Spread of Hypoxia. *Nat. Commun*., **2014**; 5: 4862 DOI: 10.1038/ncomms5862
2. Schmidt, H.; Czeschel, R.; Visbeck, M. Seasonal Variability of the Arabian Sea Intermediate Circulation and its Impact on Seasonal Changes of the Upper Oxygen Minimum Zone. *Ocean Sci.*, **2020** 16, 1459-1474, https://doi.org/10.5194/os-16-1459-2020
3. Goes, Joaquim I.; Tian, H.; do Rosario Gomes, H.; Anderson, R.; Al-Hashmi, K.; de Rada, S.; Luo, H.; Al-Kharusi, L.; Al-Azri, A.; Martinson, D. Ecosystem State Change in the Arabian Sea Fueled by the Recent Loss of Snow over the Himalayan-Tibetan Plateau Region. *Sci. Rep.*, **2020**. vol. 10, no. 1. https://doi.org/10.1038/s41598-020-64360-2
4. Scott C. Doney, Mary Ruckelshaus, J. Emmett Duffy, James P. Barry, Francis Chan, Chad A. English, Heather M. Galindo, Jacqueline M. Grebmeier, Anne B. Hollowed, Nancy Knowlton, Jeffrey Polovina, Nancy N. Rabalais, William J. Sydeman, and Lynne D. Talley. Climate Change Impacts on Marine Ecosystems Annual Review of Marine Science 2012 4:1, 11-37
5. Keeling, R. F.; Körtzinger, A.; Gruber, N. Ocean Deoxygenation in a Warming World, Annu. Rev. Mar. Sci., **2010** 2, 199–229. https://doi.org/10.1146/annurev.marine.010908.163855
6. Al Shehhi, M. R.; Gherboudj, I.; Ghedira, H. An Overview of Historical Harmful Algae Blooms Outbreaks in the Arabian Seas. *Mar. Pollut. Bull.*, **2014**. 86 (1-2), 314–324. https://doi.org/10.1016/j.marpolbul.2014.06.048
7. Goes, J.I.; do R. Gomes, H.; Al-Hashimi, K.; Buranapratheprat, A. Ecological Drivers of Green Noctiluca Blooms in Two Monsoonal-Driven Ecosystems. *GEOHAB*. **2018**, 232; 327-336. http://dx.doi.org/10.1007/978-3-319-70069-4_17
8. Ward, B.B.; Devol, A.H.; Rich, J.J.; Chang, B.X.; Bulow, S.E.; Naik, H.; Jayakumar, A. Denitrification as the Dominant Nitrogen Loss Process in the Arabian Sea. *Nature*, **2009**. 461 (7260), 78-81. https://doi.org/10.1038/nature08276
9. Hermes J.C.; Masumoto Y.; Beal L.M.; Roxy M.K.; Vialard J.; Andres M.; Annamalai H.; Behera S., D'Adamo N.; Doi T.; Feng M.; Han W.; Hardman-Mountford N.; Hendon H.; Hood R., Kido S.; Lee C.; Lee T.; Lengaigne M.; Li J.; Lumpkin R.; Navaneeth K.N.; Milligan B.; McPhaden M.J.; Ravichandran M.; Shinoda T.; Singh A.; Sloyan B.; Strutton P.G.; Subramanian A.C.; Thurston S.; Tozuka T.; Ummenhofer C.C.; Unnikrishnan A.S.; Venkatesan R.; Wang D.; Wiggert J.; Yu L.; Yu W. A Sustained Ocean Observing System in the Indian Ocean for Climate-Related Scientific Knowledge and Societal Needs. *Front. Mar. Sci*., **2019** 6, 355. https://doi.org/10.3389/fmars.2019.00355
10. Wells, M. L.; Trainer, V. L.; Smayda, T. J.; Karlson, B. S.; Trick, C. G.; Kudela, R. M.; Cochlan, W. P. Harmful Algal Blooms and Climate Change: Learning from the Past and Present to Forecast the Future. *Harmful Algae*, **2015**. 49, 68-93. https://doi.org/10.1016/j.hal.2015.07.009
11. Brooks, B.W.; Lazorchak, J.M.; Howard, M.D.A.; Johnson, M-V.V.; Morton, S.L.; Perkins, D.A.K.; Reavie, E.D.; Scott, G.I.; Smith, S.A.; Steevens, J.A. Are Harmful Algal Blooms Becoming the Greatest Inland Water Quality Threat to Public Health and Aquatic Ecosystems? *SETAC*, **2016**. V35, I1, 6-13. https://doi.org/10.1002/etc.3220
12. Laffoley, D. D. A.; Baxter, J. M.; Pitcher, G. C.; Jacinto, G. S. Ocean deoxygenation links to harmful algal blooms. In Ocean deoxygenation: Everyone's problem - causes, impacts, consequences, and solutions; IUCN, 2019; pp 153–170.
13. Turkoglu, M. Red Tides of the dinoflagellate noctiluca scintillans associated with eutrophication in the Sea of Marmara (the Dardanelles, Turkey). *Oceanologia* **2013**, 55 (3), 709–732.
14. Gomes, H. D.; McKee, K.; Mile, A.; Thandapu, S.; Al-Hashmi, K.; Jiang, X.; Goes, J. I. Influence of Light Availability and Prey Type on the Growth and Photo-Physiological Rates of the Mixotroph Noctiluca scintillans. *Front. Mar. Sci*, **2018**. 5. https://doi.org/10.3389/fmars.2018.00374
15. Lotliker, A.A.; Baliarsingh, S.K.; Trainer, V.L.; Wells, M.L.; Wilson, C.; Bhaskar, U.T.V.S.; Samanta, A.; Shahimol, S.R. Characterization of Oceanic Noctiluca Blooms Not Associated with Hypoxia in the Northeastern Arabian Sea. *Elsevier*, **2018**. 74. 46-57
16. Sarma, V.V.S.S.; Patil, J.S; Shankar, D.; Anil, A.C. Shallow Convective Mixing Promotes Massive Noctiluca scintillans Bloom in the Northeastern Arabian Sea. *Mar. Pollut. Bull.*, **2019**. 138; 428-

DOI: 10.36838/v4i6.5

436 https://doi.org/10.1016/j.marpol bul.2018.11.054
17. Poloczanska E.S., Burrows M.T.; Brown C.J.; García M. J; Halp ern B.S.; Hoegh-Guldberg O.; Kappel C.V.; Moore P.J.; Richard son A.J.; Schoeman D.S.; Sydeman W.J. Responses of Marine Organisms to Climate Change across Oceans. *Front. Mar. Sci*., **2016**. https://doi.org/10.3389/fmars.2016.00062
18. Qi, L.; Tsai, S-F.; Chen, Y.; Le, C.; Hu, C. In Search of Red Noc tiluca scintillans Blooms in the East China Sea. AGU Geophysical Research Letters, **2019**. Volume 46, Issue 11, 5997-6004. https:// doi.org/10.1029/2019GL082667
19. Liu, D.; Zhou, M. Green Tides of the Yellow Sea: Massive Free-Floating Blooms of Ulva prolifera, **2018** In Glibert P; Ber dalet E.; Burford M.; Pitcher G.; Zhou M. Global Ecology and Oceanography of Harmful Algal Blooms. Ecological Studies (Analysis and Synthesis), Springer vol 232. 317-326. https://doi. org/10.1007/978-3-319-70069-4_16
20. Suter E.A.; Lwiza K.M.M.; Rose J.M.; Gobler C.; Taylor G.T. Phytoplankton Assemblage Changes During Decadal Decreas es in Nitrogen Loadings to the Urbanized Long Island Sound Estuary, USA, **2014** Mar Ecol Prog Ser 497:51-67. https://doi. org/10.3354/meps10602
21. Doney, S.C.; Busch, S.D.; Cooley, S.R.; Kroeker, K.J. The Impacts of Ocean Acidification on Marine Ecosystems and Reliant Human Communities, **2020**. *Annu. Rev. Environ*. *Resour*. 45:1, 83-112.

Author

Alisa D'Souza graduated from Peekskill High School in 2021 with advanced designation and math and science mastery; she won 4 total awards at the Westchester Science and Engineering Fair and the Tri-County Science Fair for this research. Alisa will attend Stony Brook University, majoring in marine science.

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

CFAP418-AS1 Gene Amplification is Associated with Decreased Pancreatic Cancer Patient's Survival Rate

Daniel Rhee
Seoul Foreign School, 39 Yeonhui-ro 22-gil, Seodaemun-gu, Seoul, South Korea; gosyber1@gmail.com

ABSTRACT: Identifying specific biomarkers associated with pancreatic ductal adenocarcinoma (PDAC) patients' survival at an early stage is essential. After analyzing 1,206 pancreatic cancer patients' genomic data provided by cBioPortal database, an open web-based platform for cancer genomics, fourteen genes were found to be significantly amplified in the deceased patient group compared to the living patient group. Among those fourteen genes, it was found that *CFAP418-AS1* amplified patients (n= 9) are significantly associated with decreased median overall survival month (9.14 months) when compared with the median overall survival month (24.9 months) of *CFAP418-AS1* non-amplified patients (n= 457). To determine how *CFAP418-AS1* amplification decreased patients' survival rates, an *in vitro* assay was performed to overexpress *CFAP418-AS1* in ASPC1 pancreatic cancer cell line. In addition, Prestoblue assay indicated that overexpression of *CFAP418-AS1* increased cell proliferation. Therefore, *CFAP418-AS1* may function as an oncogene and decrease survival rates of pancreatic cancer patients by increasing cancer cell proliferation. Further investigation of the mechanisms affecting *CFAP418-AS1* dysregulation will provide insights into the molecular differences underpinning pancreatic cancer's survival rate.

KEYWORDS: Biology; Cancer Biology; Genetics, CFAP418-AS1, Gene Amplification.

Introduction

Pancreatic cancer is a relatively uncommon cancer; approximately 60,000 new diagnoses are expected in 2021 in the US.[1] However, the incidence of pancreatic cancer is increasing about 0.7% per year, and it is expected to become the second-leading cause of cancer-associated mortality by 2030.[2] The most common type of pancreatic cancer is pancreatic ductal adenocarcinoma (PDAC). This cancer is late stage, which is when pancreas cancer patients are often diagnosed.[3]

Pancreatic adenocarcinoma heterogeneity makes it difficult to predict patients' survival rates due to poor tumor cellularity and genomic instability.[4] To address this, a whole genome was analyzed from pancreatic tumors. In this research, cBioPortal, a web-based open-source cancer genomic database, was used to analyze the large-scale cancer genomics data.[5]

Lines of evidence have shown copy number variation (CNVs) of certain genes are involved in cancer progression.[6] CNV is defined as an increasing or decreasing number of DNA segments (larger than 1kb) in the human genome.[7] Currently, research focuses on somatic CNV in cancer, which provides a biological function and human disease on the genomic level. CNV is highly associated with the development and progression of many cancers by altering gene expression levels.[8]

CFAP418 Antisense RNA 1 (*CFAP418-AS1*) is a long noncoding RNA (lncRNA) gene of unknown function.[9] lncRNA regulates gene expression through interacting with nucleic acids and proteins in the cells.[10] Therefore, lncRNAs have been shown to play an important role in many different types of cancers. For example, MALAT1 is overexpressed in various cancer cells, and knockdown potently reduces both proliferation and metastasis *in vivo* in mouse model assays.[11] On the other hand, lncRNAs may also function as a tumor suppressor. p53 mediates lncRNA-p21 to induce apoptosis in cancer cells.[12]

In this study, a meta-analysis of patient data from cBioPortal database was performed to find a novel genetic marker that predicts poor outcomes in pancreatic cancer patients. The amplification of *CFAP418-AS1* was evaluated as a potential prognostic biomarker in pancreatic cancer. Herein, a possible role of *CFAP418-AS1* as a prognostic biomarker in deceased patients was discovered. Furthermore, the role of *CAFP418-AS1* by ectopic overexpression in ASPC1 (human pancreatic cancer cell line) was investigated.

Methods

Patient survival Analysis with cBioPortal:

The cBioPortal provided visualization and analyzing tools for more than 6,000 tumor samples from 290 cancer studies in TCGA database. The database provided researchers with an opportunity to analyze genetic alterations across samples from other cancer studies with specific genes. *CFAP418-AS1* was searched in cBioPortal database; a total of 1,206 samples from 10 pancreatic cancer studies were obtained. Genetic alterations such as amplifications, deep deletions, and mutations can be identified. Overall survival (OS) was calculated using cBioPortal's survival tab.

Cell line and culture:

Human pancreatic cancer cell line ASPC1 was purchased from the Korean Cell Line Bank. ASPC1 cells were cultured with RPMI-1640 medium (Gibco) supplemented with 10% fetal bovine serum (Thermo Science) and 1% penicillin and streptomycin in a 5% CO2 atmosphere at 37°C.

cDNA synthesis:

Total cellular RNA was extracted with Total RNA extraction spin kit (Intron) and treated with DNase I (Invitrogen). According to the manufacturer's protocol, the cDNA was

DOI: 10.36838/v4i6.6

reverse transcribed from 1 µg of total RNA using oligo (dT) primers (Enzynomics)

Cloning CFAP418-AS1 overexpression vector :

The cDNA of *CFAP418-AS1* was amplified with the forward primer 5′- ATAGAATTCGAGTGAAGAGGTG-CCAGAAT -3′ with EcoRI restriction site, and the reverse primer 5′- CACTCCAGCCTGGGTGACAA -3′ with XhoI restriction site. After PCR purification (Bioneer), the amplified product was digested with EcoRI and XhoI restriction enzyme. After PCR purification, the amplified DNA was cloned into pcDNA3 (addgene) vector.

Cell transfection :

Lipofectamine 2000 (Invitrogen) was used to transfect *pcDNA3-CFAP418-AS1* in ASPC1 cell line. Diluted Lipofectamine to the diluted DNA (1:1 ratio) was used to transfect the cells. After incubating for 15 min at room temperature, DNA-lipid complexes were added to the cultured cells drop-wise. The cell culture media was replaced after three hours. Downstream experiments were performed 48 hours post-transfection.

Polymerase Chain Reaction :

20 µL of Polymerase Chain Reaction (PCR) was performed with PCR-Premix (Bioneer). The primers were designed to amplify CAFP418-AS1 and GAPDH. The cycles of 94˚C for 15 s, 55˚C to 60˚C for 15 s, and 74˚C for 30 s were used to amplify the target gene. The forward primer 5′- TGGCCATGAGGGATTCAAGG -3′ and reverse primer 5′- GAACACACTGTGCTGTCCCT -3′ yielded a 172-base pair (bp) product from CFAP418-AS1 using an annealing temperature 60˚C. For GAPDH amplification, 176 bp of amplified product was synthesized using forward primer 5′- TGGAGAAGGCTGGGGCTCAT -3′ and reverse primer 5′- GACCTTGGCCAGGGGTGCTA -3′.

Cell transfection :

RedSafe nucleic acid staining reagent (Intron) was used to stain 1.3% agarose gels. After samples were loaded, agarose gels were run at 100 volts for 25 min. The gel electrophoresis images were captured with a digital camera, and the Image J program was used to quantify band intensities. The relative intensities of *CFAP418-AS1* bands were normalized by the GAPDH band.

Statistical analysis :

All statistical analysis was performed using Prism 7 program. Unpaired t-test was used to calculate the statistical significance. $p < 0.05$ was considered statistically significant.

Results and Discussion

Table 1: The list of amplified genes that are enriched in deceased pancreatic patient's group.

Gene	Cytoband	Number of living patients (%)	Number of decreased patients (%)	Log Ratio	p-Value	q-Value	Enriched in
TP53	17p13.1	88 (34.92%)	128 (59.81%)	-0.78	5.83E-08	1.226E-03	DECEASED
CFAP418-AS	8q22.1	0 (0.00%)	9 (9.09%)	<-10	8.614E-06	0.0325	DECEASED
LINC02894	8q22.1	0 (0.00%)	9 (9.09%)	<-10	8.614E-06	0.0325	DECEASED
DHX36	3q25.2	0 (0.00%)	13 (6.53%)	<-10	1.917E-05	0.0325	DECEASED
DNAI1	9p13.3	0 (0.00%)	13 (6.53%)	<-10	1.917E-05	0.0325	DECEASED
PMEPA1	20q13.31	0 (0.00%)	13 (6.53%)	<-10	1.917E-05	0.0325	DECEASED
C8ORF34-AS1	8q13.2	0 (0.00%)	8 (8.08%)	<-10	3.247E-05	0.0325	DECEASED
C8ORF88	8q21.3	0 (0.00%)	8 (8.08%)	<-10	3.247E-05	0.0325	DECEASED
C8ORF89	8q21.11	0 (0.00%)	8 (8.08%)	<-10	3.247E-05	0.0325	DECEASED
LINC01030	8q21.3	0 (0.00%)	8 (8.08%)	<-10	3.247E-05	0.0325	DECEASED
LINC01111	8q21.13	0 (0.00%)	8 (8.08%)	<-10	3.247E-05	0.0325	DECEASED
LINC01298	8q22.1	0 (0.00%)	8 (8.08%)	<-10	3.247E-05	0.0325	DECEASED
LINC01592	8q13.2	0 (0.00%)	8 (8.08%)	<-10	3.247E-05	0.0325	DECEASED
MIR2052HG	8q21.11-q21.13	0 (0.00%)	8 (8.08%)	<-10	3.247E-05	0.0325	DECEASED

Meta-analysis of 1,206 pancreatic cancer patients' genomic data provided by cBioPortal database, 14 genes were found to be significantly amplified in deceased patient group compared to living patient group (Table 1). Among 14 genes, ten genes were located on chromosome 8. Since amplification of CFAP418-AS was analyzed as one of the top significant alterations enriched in deceased pancreatic patient group, CFAP418-AS was the focus of amplification for downstream analysis.

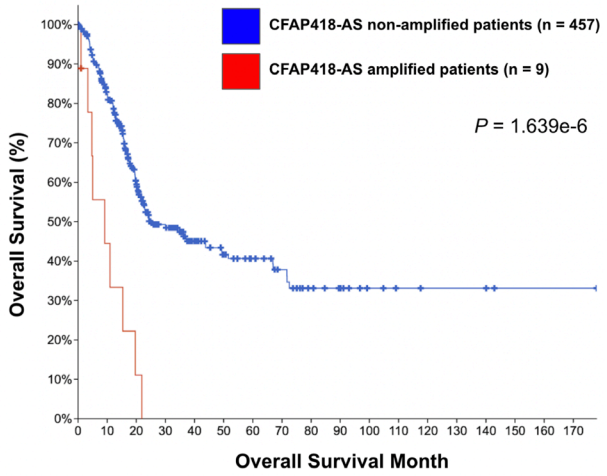


Figure 1: Kaplan-Meier survival plot of patients with pancreatic cancer, assessing overall survival percentage in regard to CFAP418-AS amplification. The patient samples were split into two groups according to CFAP418-AS amplified status (n= 9) and non-amplified (n= 457). The overall survival curve of gastric cancer patients between two groups was analyzed by Kaplan–Meier survival plot (p=1.639e-6).

To determine if the amplification of CFAP418 can be predictive for pancreatic cancer patient survival, 466 patient samples were analyzed using cBioPortal to validate the gastric cancer survival biomarker candidates. The patient samples were divided into two groups according to the amplification status of CFAP418-AS. After performing the patient survival analysis using Kaplan-Meier Plot, two patient cohorts (amplified vs. non-amplified of CFAP418-AS) were compared. As expected, patients with CFAP418-AS amplification showed a statistically significant decrease in overall survival (p = 1.639e-6) (Figure 1). Overall, CFAP418-AS amplification is significantly associated with a low survival rate of pancreatic patients.

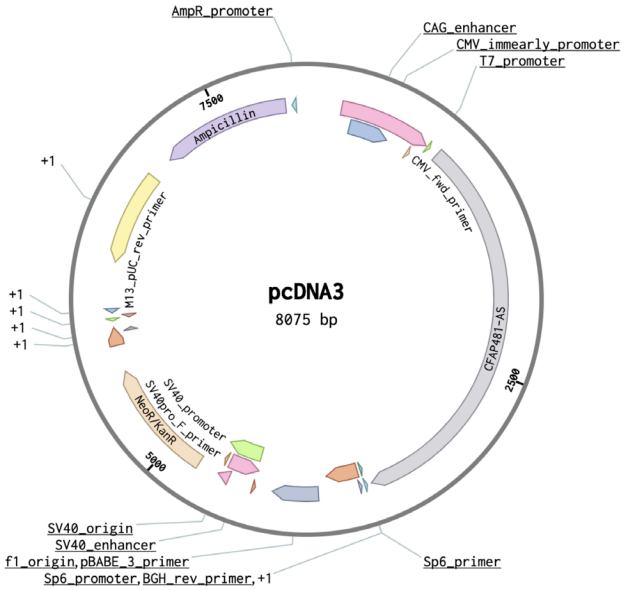


Figure 2: pcDNA3 plasmid with CFAP418-AS gene (pcDNA3-CFAP418-AS).

CFAP418-AS gene (2659bp) was cloned into pcDNA3 plasmid to investigate the functional role of CFAP418-AS gene in pancreatic cancer. pcDNA3 is a mammalian expression vector with the CMV promoter. It is a widely used vector for overexpressing the specific gene of interest. The human gene CFAP418-AS was successfully cloned into pcDNA3 plasmid (Figure 2).

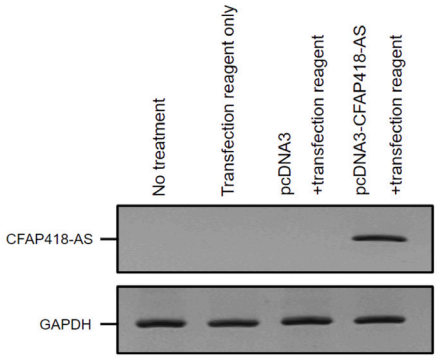


Figure 3: pcDNA3-CFAP418-AS transfected ASPC1 cells overexpress CFAP418. The amplified cDNA from each sample by PCR was analyzed with 1.3% agarose electrophoresis gel.

Four different conditions were used to verify the overexpression of CFAP418-AS gene on ASPC1 cells: no treatment (negative control), transfection reagent only, pcDNA3 + transfection reagent, pcDNA3-CFAP418-AS + transfection reagent. Overexpression of CFAP418 was expected only in pcDNA3-CFAP418-AS transfected samples. After DNA transfection on ASPC1 cells, RNA was extracted, and cDNA was synthesized by RT-PCR. cDNA was amplified with specific primer pairs that target CFAP418-AS and GAPDH genes by PCR. Then agarose gel electrophoresis was performed to check the amplified DNA. Agarose gel data indicates that CFAP418-AS was amplified only in pcDNA3-CFAP418-AS transfected cells showing a band on CFAP418-AS (Figure 3). GAPDH band showed a similar intensity of bands on all four conditions, indicating a similar total RNA and cDNA was used for each condition. Overall, pcDNA3-CFAP418-AS successfully overexpressed CFAP418-AS.

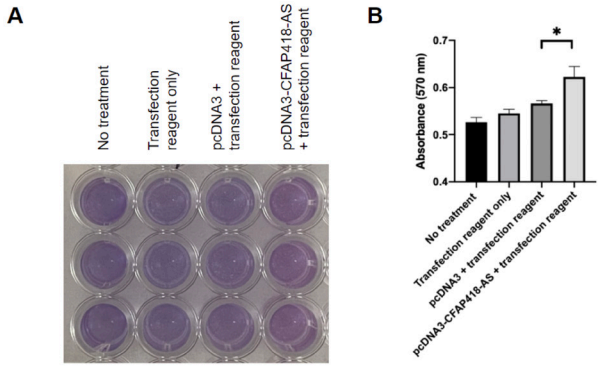


Figure 4: CFAP418-AS overexpression increased cell proliferation in the pancreatic cancer cell line, ASPC1. (A) Comparison of cell proliferation using Prestoblue reagent (B) Quantification of cancer cell proliferation based on the 570nm absorbance measurement. The number of samples = 3, Student's t-test. (*, p< 0.05)

Cell proliferation is how quickly a cancer cell replicates its DNA and divides into two cells.[13] Therefore, an increase in cell proliferation is associated with faster-growing tumors and is more aggressive. Prestoblue assay was performed to investigate the effect of overexpression of CAFAP418-AS on pancreatic cancer cell proliferation. pcDNA3-CFAP418-AS transfected cells showed the highest cancer cell proliferation (Figure 4). When the proliferation of pcDNA3 transfected cells was compared to the pcDNA3-CFAP418-AS transfected cells, the cell proliferation increased to about 120%. This result indicated that CFAP418-AS overexpression may positively regulate pancreatic cancer proliferation.

Conclusion

Antisense gene expression can be regulated either coordinately or independently of their neighboring genes.[14] Also, antisense transcripts can regulate the expression of their target genes from transcription and translation to RNA degradation.[15] CFAP418-AS gene is an antisense gene that may regulate CFAP418 gene expression level. CFAP418 encodes a protein of unknown function. CFAP418 is ubiquitously expressed in many organs such as the brain, heart, retina, and pancreas.[9] This study found that amplification of CFAP418-AS gene is significantly associated with decreased pancreatic patients' survival. Also, this *in vitro* experiment showed that overexpression of CFAP418-AS may enhance pancreatic cancer progression. In conclusion, this study indicates that CFAP418 amplification or overexpression can be used as a biomarker that predicts poor prognosis in pancreatic cancer. Also, this study can be used to develop a novel cancer treatment. However, further study is needed to investigate the detailed molecular mechanism of how CFAP418-AS functions as an oncogene in pancreatic cancer.

Acknowledgements

I would like to give a special thanks to Dr. Woo Rin Lee from University of Suwon for his guidance on this project.

References

1. Park, W.; Chawla, A.; O'Reilly, E. M. Pancreatic Cancer: A Review. *JAMA* **2021**, 326 (9), 851–862.

DOI: 10.36838/v4i6.6

2. Huang, J.; Lok, V.; Ngai, C. H.; Zhang, L.; Yuan, J.; Lao, X. Q.; Ng, K.; Chong, C.; Zheng, Z.-J.; Wong, M. C. S. Worldwide Burden of, Risk Factors for, and Trends in Pancreatic Cancer. *Gastroenterolo gy* **2021**, 160 (3), 744–754.
3. McGuigan, A.; Kelly, P.; Turkington, R. C.; Jones, C.; Coleman, H. G.; McCain, R. S. Pancreatic Cancer: A Review of Clinical Diagnosis, Epidemiology, Treatment and Outcomes. *World J. Gas troenterol*. **2018**, 24 (43), 4846–4861.
4. Dell'Aquila, E.; Fulgenzi, C. A. M.; Minelli, A.; Citarella, F.; Stel lato, M.; Pantano, F.; Russano, M.; Cursano, M. C.; Napolitano, A.; Zeppola, T.; et al. Prognostic and Predictive Factors in Pancreatic Cancer. *Oncotarget* **2020**, 11 (10), 924–941.
5. Gao, J.; Aksoy, B. A.; Dogrusoz, U.; Dresdner, G.; Gross, B.; Sumer, S. O.; Sun, Y.; Jacobsen, A.; Sinha, R.; Larsson, E.; et al. Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the CBioPortal. *Sci. Signal*. **2013**, 6 (269), pl1.
6. Shlien, A.; Malkin, D. Copy Number Variations and Cancer. *Ge nome Med*. **2009,** 1 (6), 62.
7. Liu, B.; Morrison, C. D.; Johnson, C. S.; Trump, D. L.; Qin, M.; Conroy, J. C.; Wang, J.; Liu, S. Computational Methods for De tecting Copy Number Variations in Cancer Genome Using next Generation Sequencing: Principles and Challenges. *Oncotarget* **2013**, 4 (11), 1868–1881.
8. Shao, X.; Lv, N.; Liao, J.; Long, J.; Xue, R.; Ai, N.; Xu, D.; Fan, X. Copy Number Variation Is Highly Correlated with Differen tial Gene Expression: A Pan-Cancer Study. *BMC Med. Genet*. **2019**, 20 (1), 175.
9. Safran, M.; Dalah, I.; Alexander, J.; Rosen, N.; Iny Stein, T.; Shmoish, M.; Nativ, N.; Bahir, I.; Doniger, T.; Krug, H.; et al. GeneCards Version 3: The Human Gene Integrator. *Database (Ox ford)* **2010**, 2010, baq020.
10. Jarroux, J.; Morillon, A.; Pinskaya, M. History, Discovery, and Classification of LncRNAs. *Adv. Exp. Med. Biol.* **2017**, 1008, 1–46.
11. Gutschner, T.; Hämmerle, M.; Eissmann, M.; Hsu, J.; Kim, Y.; Hung, G.; Revenko, A.; Arun, G.; Stentrup, M.; Gross, M.; et al. The Noncoding RNA MALAT1 Is a Critical Regulator of the Me tastasis Phenotype of Lung Cancer Cells. *Cancer Res*. **2013**, 73 (3), 1180–1189.
12. Yang, T.; Zhang, W.; Wang, L.; Xiao, C.; Guo, B.; Gong, Y.; Liang, X.; Huang, D.; Li, Q.; Nan, Y.; et al. Long Intergenic Noncoding RNA-P21 Inhibits Apoptosis by Decreasing PUMA Expression in Non-Small Cell Lung Cancer. *J. Int. Med. Res*. **2019**, 47 (1), 481–493.
13. Guo, M.; Hay, B. A. Cell Proliferation and Apoptosis. *Curr. Opin. Cell Biol*. **1999**, 11 (6), 745–752.
14. Brophy, J. A. N.; Voigt, C. A. Antisense Transcription as a Tool to Tune Gene Expression. *Mol. Syst. Biol*. **2016**, 12 (1), 854.
15. Pelechano, V.; Steinmetz, L. M. Gene Regulation by Antisense Transcription. *Nat. Rev. Genet*. **2013**, 14 (12), 880–893.

Author

Daniel Rhee is a high school student who currently attends Seoul Foreign School located in Seoul, South Korea. He is a swimmer, a competitor, and a STEM student. He shows deep passion in the idiosyncrasies he takes part in and always strives to be first in all that he does. In the future, he wishes to go on to study biological sciences at a decorated university.

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

The Differential Analysis and Modeling of Lung Cancer Survival Rates

Melissa Erdem
Paul J. Hagerty High School, 3225 Lockwood Blvd, Oviedo, FL, 32765,USA ; melissaerdem0@gmail.com

ABSTRACT: Lung cancers are the most prevalent cancer type with the lowest 5-year survival rate. Biases in lung cancer screening contribute to the inaccurate estimation of patient surviving days, demanding an objective model of survival estimation. This paper offers a novel two-phase approach to the modeling of lung cancer patient survival featuring a newly established VB-based Cancer Data Analysis Software Application. Given patient data gathered from the NIH's SEER Program, the author-developed software first analyzed the impacts of 24 individual genetic, diagnostic, and treatment variables on survival. In the next phase, the software successfully integrated 7 diagnostic and 6 treatment-related factors to estimate diagnostic and treatment effectiveness factors- representing the weights of various treatment and diagnostic factors on surviving days. In the final phase, coefficients were tested on sample sizes of 12,000 to 20,000 respiratory cancer patients for survival estimation. Percent error between estimated and actual survival was calculated, along with Spearman Rank correlations for significance testing. Results indicated that the software was most reliable for the following diagnostic factors: Metastases in the Lung, Metastases in Bone, number of Positive Lymph Nodes, and the control (No diagnostic variables). This VB-based software can be utilized to determine the ideal treatment combination that provides the most accurate diagnosis of surviving days by diagnostic variable. Through consideration of strict numerical data and complexity in treatment, results are ensured to be representative of the patient's true situation and not error-prone subjective observations.

KEYWORDS: Biomedical and Health Sciences; Other; Oncology; Lung cancer survival; Mathematical Modeling.

Introduction

Lung Cancer Prevalence:

Lung cancer is the deadliest and third most common cancer in the United States.[1] The National Institutes of Health report that roughly 22% of all cancer-related deaths in 2021 were attributed to the lung and bronchus, the highest percentage of deaths reported by any single cancer type (Figure 1).[1] Alarmingly, lung cancers have displayed a 22% five-year survival rate, extremely low compared to more optimistic rates such as that of breast cancer (90%).[2] Elderly individuals typically possess this cancer, and treatments usually involve combinations of chemotherapy, radiation therapy, surgery, and targeted therapy drugs, among other treatments.[3] Lung cancer's widespread prevalence and alarming survival rate demand a need for further research into the existing prognosis models for respiratory cancer patients and the exploration of individualized treatment options to maximize patient survival time.

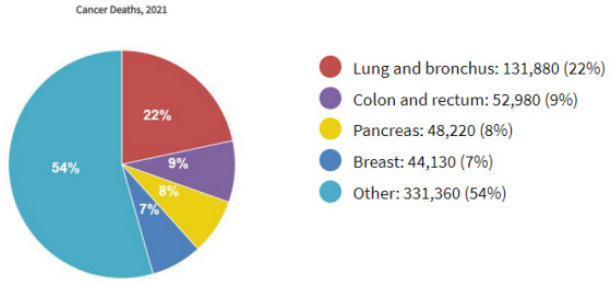


Figure 1: This figure from the National Cancer Institute depicts the percentages of total cancer deaths in the U.S. attributed to each cancer type in 2021, with respiratory cancers represented by the red section.

Pitfalls of TNM Staging:

Currently, the eighth edition of the American Joint Committee on Cancer's (AJCC) TNM system is widely used among healthcare practitioners to predict the stage of malignant tumors, focusing on the importance of tumor size, lymph node involvement, and metastases in prognosis.[4] This system is summarized in Figure 2.

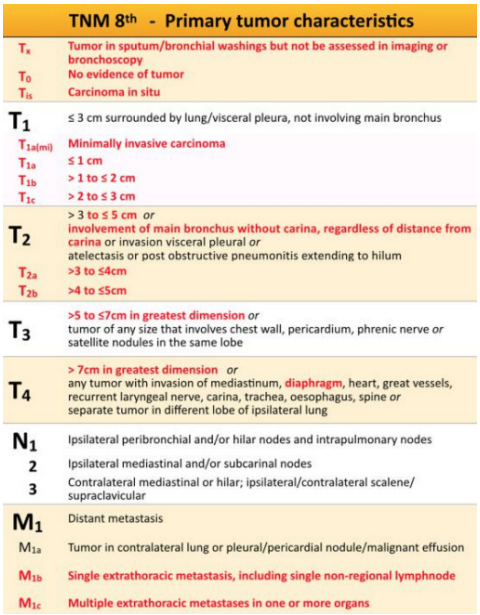


Figure 2: This figure provides an overview of the 8th edition of the Tumor, Metastases, and Lymph Node (TNM) System used for lung cancer diagnosis and established by the AJCC; retrieved from Radiology Assistant.[5]

DOI: 10.36838/v4i6.7

The AJCC highlighted numerous changes to the TNM system to promote accurate prognosis in its 8th edition, including dividing the T1 category into 3 subcategories (T1A-T1C) and adjusting and subdividing tumor sizes in the T2, T3, and T4 categories.[6] These efforts were successful, as supported by a study published in the Journal of Thoracic Disease. Yun *et al.* concluded that the latest edition had increased prognostic validity compared to its predecessor, the 7th edition, due to stratification.[7] However, studies such as Hattori *et al.*'s note persistent barriers including clinicopathological issues in the T category, and the overall lack of consideration of prognostic indicators at the molecular level, specifically driver oncogene and immune status in the evaluation of the system's applications in Non-Small Lung cancer prognosis.[8] Further, Zurrida, and Veronesi proposed the need for updated definitions for the T, N, and M categories as well as placeholders to specify the molecular characteristics of the tumor to incorporate more precise pathological information.[9] Thus, the persistent shortcomings of the TNM system and lack of histological indicators demand supplement tools for a reliable prognostic system to be established for lung cancer patients.

Paired with the lack of individualized prognosis, biases in the lung cancer screening process also contribute to inaccurate survival estimation. In their comparative study, Ge *et al.* found evidence of lead-time bias in lung cancer screening in the form of overestimated survival times for patients who previously had cancer although they posed no clinical advantage.[10] Lead-time biases in this context refer to the detection of lung cancer before the emergence of symptoms during the cancer screening stage, though the course of the disease is unaltered. As a result, the extreme increase in 5-year survival rates due to earlier diagnosis when survival is unaffected is an illusion contributing to this form of bias. Thus, the possession of no "clinical advantage" in the context of Ge *et al.*'s study is interpreted as no relative increase in surviving days. Additionally, length biases were noted in an article for the Journal of the American College of Radiology, where Gill *et al.* proposed the possibility for lung cancer screening to underestimate survival based on relatively harmless tumors posing no threat to survival.[11] Despite the persistence of over-and under-diagnosis ultimately leading to overtreatment, many patients are still not informed of these risks. A study by Wegwarth *et al.* discovered that fewer than 10% of lung cancer patients in the study responded that their doctor informed them of respective risks of overdiagnosis and/or overtreatment.[12] Moreover, due to the TNM system's non-holistic criteria for the classification of tumor stage and screening biases, health practitioners tend to either overestimate or underestimate lung cancer patient survival. Lambden et al. discovered that physicians' estimates favored overestimations of survival much more frequently as the patient's death came nearer by a median of 4.4 months when the patients had 0-3 months left to live.[13] On the contrary, physicians underestimated survival by a median of 12 months when the patient survived longer (>12 months).[13] Consequently, there exists a need to objectively predict lung cancer survival to the maximum accuracy through consideration of individual diagnostic factors and treatments in supplementation to predetermined prognosis models

Existing Studies and Limitations:

Now more prevalent than ever, Machine and Deep learning models are at the forefront of studies aiming to model lung cancer survival through consideration of diverse patient variables. A study by Lynch et al. utilized Supervised Machine Learning Classification techniques to predict lung cancer survival, incorporating Linear Regression, Decision Trees, Gradient Boosting Machines, Support Vector Machines, and more.[14] Another study published in 2020 implemented Deep Learning for the development of a model used to determine non-small cell lung cancer survival among roughly 18,000 patients.[15] Further, a study published in the International Journal of Medical Informatics in February of 2021 was able to predict lung cancer survival with 71% accuracy utilizing Deep Learning, outperforming other models.[16] Other studies have maintained the use of traditional regression models such as the Cox proportional hazards and the Kaplan Meier models in their statistical analyses. However, little to of such studies have employed a traditional mathematical modeling structure for cancer survival estimation. Typically, mathematical modeling is viewed as an "outdated" approach to the complex data analyses that cancer diagnoses warrant. However, mathematical models are complementary to newer Machine Learning approaches as they address the latter's key limitations. In the case of Lynch *et al.*'s study, survival was only estimated accurately for patients who lived for 6 months or less with models struggling to estimate survival for patients living 35 months or more.[14] Moreover, the standard error was inflated in this study in question due to large errors. Most importantly, Lynch *et al.* state that the Machine Learning techniques studied were similar in performance to the Cox proportional hazards models, implying that traditional methods should not be overlooked and should rather be used in collaboration with newer models. Studies with ML also lack proper data interpretation, which is feasible through mathematical modeling. Comparing the use of this approach to newer deep learning models can help further the understanding of why specific variables impact survival more than others, which is otherwise undiscovered with ML/DL. Another key limitation of the current literature is within the analysis of treatment information. Most studies examined did not consider multiple treatment combinations, or the diversity of treatment options a patient may undergo in each period. This overlooks the actual situation of many cancer patients, as authors assume that patients undergo a single treatment such as Chemotherapy alongside a diagnostic or social variable. Thus, biases are introduced when not accounting for the percentage of patients who undergo multiple treatments simultaneously.

Methods

Research Design:

This study employed a purely quantitative approach featuring an applied multivariate analysis implementing matrix algebra to estimate treatment and diagnostic variable effectiveness factors, as well as subsequent surviving days. For the data

that may have prevented a linear association between variables and surviving days. Further, to be included in each analysis, patients needed to have non-null values for all six treatments and the variable in question. Despite data collection spanning the years 1975-2017, only data from 2004 and later was utilized to ensure accurate collection mechanisms and to include variables with the prefix "CS" (Stage-related variables) that were only collected starting from the year 2004. Overall, nearly 3 million cancer patients were filtered to sample sizes of 25-100K patients (which varied among diagnostic variables). As a critical part of the procedure, mean values of surviving days were taken for patients who underwent similar treatments. This served as an attempt to normalize the patient data to avoid the possibility of non-linear phenomena (i.e., two patients with similar treatment options and drastically different surviving days). Sample sizes of patients were randomly selected through a module developed in the SQL Server. A database in SQL labeled "Respiratory Cancer Data" was thus created, storing the filtered and reformatted variable data, de-identified patient IDs, and actual surviving days for all patients (reformatted from the original month format). Moreover, Visual Basic was utilized to implement the Cancer Data Analysis Control Panel display's back-end code for loading and analyzing data. Primary programming languages included the Visual Basic Language, Java, and MATLAB. Modules were developed to evaluate single variable effects on survival, followed by software-generated box plots displaying means of surviving days for each variable's possible values (see Figure 7). This single-variable analysis was followed by a differential analysis segment implemented to calculate diagnostic and treatment effectiveness factors provided for each of the seven diagnostic variables. For this, a mathematical model was implemented as seen in Figure 3. For each of the seven diagnostic variables, the mathematical equation below was implemented to obtain treatment and diagnostic coefficients (stored in Vector X) by multiplying the inverse of the patient data (represented by Matrix A) by the actual number of surviving days for each patient (represented by vector B_{actual}). Specifically, Matrix A stored the two-dimensional patient data (diagnostic and treatment information in the columns and patient IDs in the rows). To obtain the inverse of Matrix A, a pseudo-inverse was taken using a third-party Visual Studio plugin. The ultimate objective of estimating the coefficients was to later implement the calculated treatment and diagnostic effectiveness factors in survival estimation.

$f(x)$= Surviving days = $\sum_{i=1}^{7} a_i \, . \, x_i$ where x_1 is the selected diagnostic variable's coefficient and x_2-x_7 are 6 respective treatment coefficients.

$x = A^{-1} . b_{actual}$ yielded estimated *diagnostic and treatment coefficients*

$$\begin{bmatrix} x_1 \\ x_2 \\ \\ x_7 \end{bmatrix} = A^{-1} \times \begin{bmatrix} b_{act_1} \\ b_{act_2} \\ \\ b_{act_n} \end{bmatrix}$$

Figure 3: This figure displays the mathematical modeling structure implemented in the "Differential Analysis" component of Phase 1 of the study. On the left-hand side of the equation Vector X represents the calculated diagnostic and treatment coefficients whereas on the right-hand side Matrix A represents patient data and Vector Bactual represents actual surviving days of patients.

Supplementing the differential analysis, modules were written to display bar charts visualizing the calculated coefficients following the pop-up of text (see Figure 8). In the survival estimation phase of software development, new functions were developed in the Visual Basic platform. Seven diagnostic variables were chosen for analysis due to their primary role in cancer prognosis under the TNM System. Similar to the first phase, a control group was utilized where no diagnostic variable was involved, and the same six treatment variables persisted from the first phase. Moreover, calculated treatment effectiveness factors were obtained and loaded into a new data table using the Microsoft SQL server and presented in Vector X. This time, to-be estimated surviving days were stored in a $b_{estimate}$-Vector. Unlike in Phase 1, the pool of patients was reduced to 10,000-20,000 patients depending on data availability for each diagnostic variable since mean surviving days were taken from patients with similar treatment combinations as opposed to similar individual treatments. This choice was undertaken to eventually determine which combinations could provide the most accurate surviving days. Hence, the normalization of matrices based on a non-linear system allowed for the manipulation of the same mathematical equation for more accurate survival estimation as demonstrated in Figure 4.

$$A * x = b_{estimate}$$

Matrix Algebra after Normalization:

$$\begin{bmatrix} a_{11} & a_{12\ldots} & \ldots & \ldots & a_{17} \\ a_{21} & a_{22\ldots} & \ldots & \ldots & a_{27} \\ & & & & \\ a_{n1} & a_{n2} & \ldots & \ldots & a_{n7} \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ \\ x_7 \end{bmatrix} = \begin{bmatrix} b_{est_1} \\ b_{est_2} \\ \\ b_{est_n} \end{bmatrix}$$

*In the case of the Metastases at DX Bone variable survival analysis, n= 13, indicating 13 unique treatment combinations.

Figure 4: This figure displays the matrix algebra unique to this year's segment in estimating survival utilizing previously calculated diagnostic and treatment effectiveness factors (in the x-Vector).

Back-end code was then implemented for the Cancer Data Analysis Software to develop bar graphs following each variable analysis, displaying the treatment combinations on the x-axis and the survival estimation error (in days) on the y-axis. The sample size of that variable was noted in a legend at the top right. These graphs were developed to display which treatment combinations produced the least estimation error. Percent error and median calculations were then conducted between estimated and actual surviving days. To determine the significance between estimated and actual survival, the Spearman Ranks Correlation test was utilized. This non-parametric test was chosen because of the traditionally non-normal distribution of data and is most like the parametric Pearson Correlation test.

DOI: 10.36838/v4i6.7

Materials and Risk Assessment:

To conduct this study, lung cancer patient data were retrieved from the National Institute of Health's Surveillance, Epidemiology, and End Results Program. The program contains population and incidence cancer information for a variety of cancer types and includes 70+ social, diagnostic, and treatment variables. After registering through a cancer researcher SEER account and signing appropriate confidentiality and risk assessments, raw patient data was accessible. However, treatment information was only obtained through signing an additional form for access to SEER's specialized treatment database. By signing the SEER forms acknowledging patient confidentiality and safety measures, the ethicality of this study was ensured. Furthermore, access to a computer to develop the Cancer Data Analysis Control Panel with Visual Basic Studio and Microsoft SQL to analyze the raw data was also necessary. To maintain safety, the 20-20-20 rule for safe computer use was followed. No further risks were presented in this study.

Seer Program and Participant Information:

TRepresentative of the United States' population diagnosed with cancer, the NIH's Surveillance, Epidemiology, and End Results Program collects cancer survival data from approximately 35% of the U.S population and mimics the racial and socioeconomic diversity of the nation.[19] The 19 registries nationwide are comparable to the larger U.S. population in poverty and education levels, with percentages of SEER patients with less than a high school diploma and those below poverty similar to those of the general population.[19] However, SEER participants contain a slightly greater percentage of foreign-born individuals compared to the general population.[19] The program has since expanded its registries for coverage of Hispanics, urban African Americans, Asian and Pacific Islanders in Southern California and the Greater Bay Area, rural African Americans in Georgia, Arizona Indians, and Alaska natives.19 Key limitations of other cancer patient databases such as the Society of General Internal Medicine's National Cancer Database are that cohorts are not population-based; rather, they are identified from hospitals.[20] On the contrary, SEER data is known as the "gold standard in data quality" among most U.S. registries and holds itself to strict data quality and collection policies.[21] By using the most representative data for cancer analysis, it can be ensured that conclusions apply to all respiratory cancer patients from across diverse regions in the United States.

Variables and Hypothesis:

The independent variables in the first phase of the study included 24 various social, diagnostic, and treatment variables, including actual surviving days. The four social variables listed were selected due to having the most complete data for the time frame of analysis: marital status, age at diagnosis, race, and gender. Thirteen diagnostic variables selected under the same reasoning included Metastases at Diagnosis, Primary Site, Laterality, Number of Positive Lymph Nodes, Tumor Size, and more. Moreover, six treatment variables included Surgery conducted at Primary Site, Chemotherapy, Radiation therapy, Intraoperative Radiation , Radiation before Surgery, and Radiation after Surgery. These therapies were the only treatment-related variables available in the specialized Radiation/Chemotherapy database, with data for newer therapies like immunotherapy or targeted therapies limited in the scope of analysis, and thus unused. The control group was the analysis of patients without any diagnostic variable, providing a baseline for comparisons among variables. In the differential analysis component of the first phase and survival estimation in the second phase, seven diagnostic variables and six treatment variables were analyzed, including control of no diagnostic variables. Diagnostic variables included Positive Lymph Nodes, Metastases in Lung, Metastases at Bone, Metastases in all Organs, Primary Site, Tumor Size, and Total Number of In-Situ tumors. The advantage posed by focusing on a limited scope of diagnostic variables included a more thorough and reasonable analysis of significance in the relationship between estimated and actual survival for each variable (quality over quantity). Moreover, diagnostic variables were continuous (as in 3 mm for tumor size), whereas treatment variables were discrete (as in Yes/No for chemotherapy). Consequently, whether or not the patient received treatment or not as indicated by either a "0" or "1". Actual survival was measured in days, converted from the original month form provided by SEER. Within the single-variable and differential analyses in Phase 1, the dependent variable was the number of actual surviving days. As for survival estimation in Phase 2, the dependent variable was the estimated number of surviving days after matrix multiplication. Moreover, attrition and confounding variables were risks to the internal validity of this study, since it was not guaranteed by SEER that all data collected was accurately measured. In the first phase (including single-variable and differential analyses), it was hypothesized that the number of surviving days of a respiratory cancer patient was dependent upon the social, genetic, and diagnostic variables and that the normalized mathematical modeling system would produce a linear relationship between surviving days and variable quantities. In the second phase, it was predicted that given effectiveness factors, there would be a near-zero percent change between estimated and actual survival and that the relationship between the two measures would be significant for all diagnostic variables.

Data Analysis:

Percent error calculations were conducted between estimated and actual survival data for each variable, along with mean calculations of the estimated surviving days for each variable. Patterns in mean calculations among diagnostic variables were later compared to the results of prior studies in the field.

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)}$$

ρ = Spearman's rank correlation coefficient
d_i = difference between the two ranks of each observation
n = number of observations

Figure 5: This figure depicts the equation for the calculation of a Spearman Ranks Coefficient (noted as ρ or *rs*), retrieved from Statistics How To.

[1] According to the Mayo Clinic, Intraoperative Radiation therapy (IORT) is an intensive radiation treatment administered during surgery, providing direct radiation to a specific area.

Results

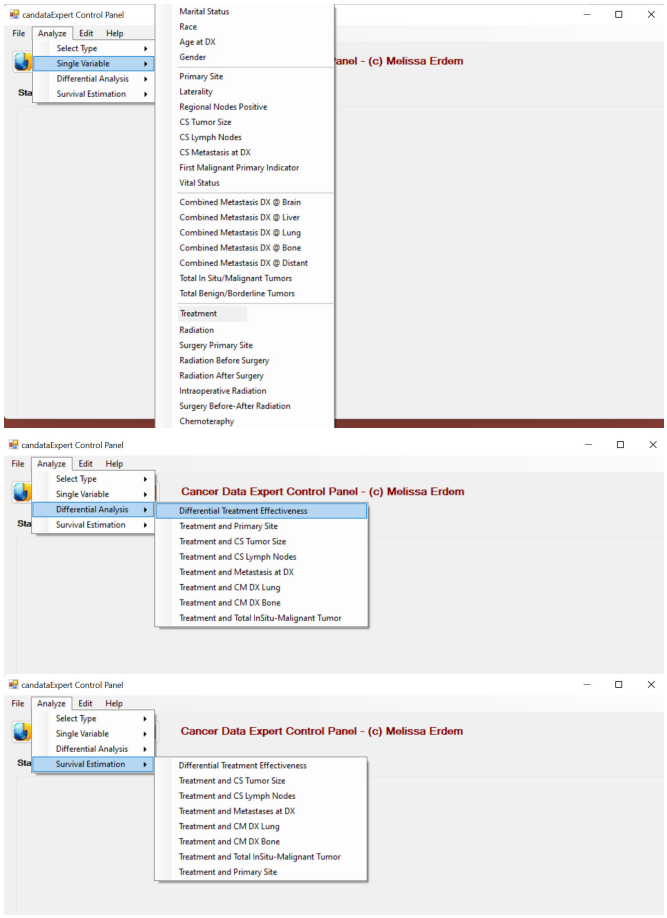
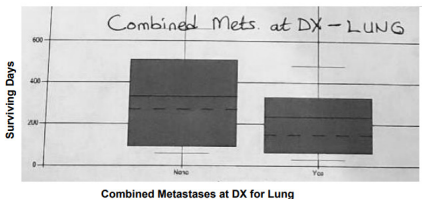


Figure 6: Screenshots of the Cancer Analysis Software Control Panel, demonstrating where to navigate to run a single-variable analysis (top image), a differential analysis (middle image), and a survival estimation (bottom image).

Data Collection:



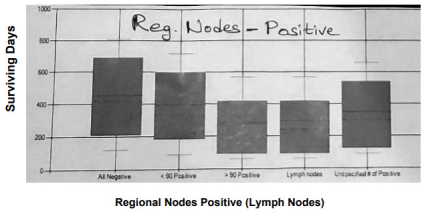


Figure 7: The first two graphs (a.) and (b.) are box plots (with mean and median actual surviving days) displayed by the Cancer Data Analysis Software under the single-variable analysis of the first phase. Only graphs for two diagnostic variables are featured (Combined Metastases at Diagnosis for Lung and Positive Lymph Nodes respectively). The final graph (c.) shows a multiple-variable analysis under the first phase, where six treatments are analyzed in respect to the Metastases at Diagnosis variable.

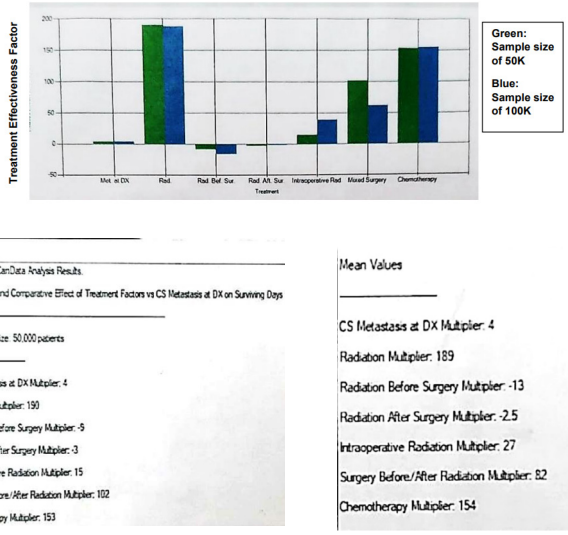


Figure 8: The following is a screenshot of the software's display when a differential analysis is conducted with the Metastases at Diagnosis variable. The images depict a graph and resulting status box containing sample and mean diagnostic and treatment coefficients.

Table 1: This chart organizes calculated treatment effectiveness factors by diagnostic variable (columns) and treatments (rows) as a result of the differential analysis portion of Phase 1 of the study.

Mean Diagnostic and Treatment Coefficients of 25-100K Patients by Diagnostic Variable

Treatments	Control (No Diagnostic Variables)	CS Metastases at Diagnosis	CS Lymph Nodes	CS Tumor Size	Primary Site Malignant Tumors	Metastases at Lung	Metastases at Bone
Diagnostic Variable Coefficient	N/A	4	0	0	1	2	1	0
Radiation	183.67	189	167	163.67	7.5	85.4	198	180
Radiation Before Surgery	-20	-13	-18.5	-2	40	21	15	-20
Radiation After Surgery	-10	-2.5	-21	-0.67	19	-10	24	50
Intraoperative Radiation	42.67	27	-7.5	18.67	35.5	-2.5	10	3
Surgery Before & After Surgery	-25	82	-38	-31.33	-1.5	20	-2	-10
Chemotherapy	207.67	154	154	174	27	93.5	180	170

After running the software, diagnostic and treatment coefficients shown in Table 1 were obtained, representing the weights of each diagnostic variable and six other treatments on surviving days. Interpretation of the meaning of the integers representing the coefficients was not possible at this stage of the study, solely because the calculation of coefficients was an intermediate step in the ultimate goal of estimating survival using the coefficients. As seen in the chart, there were instances of negative coefficients for certain diagnostic/treatment pairs. As such, interpreting the coefficients as "more impactful" if they were greater would not be logical. Hence, the presence of negative coefficients inexplainable in mathematics proves the non-linearity of the system.

In the second phase, lung cancer treatment combinations for a total of 7 diagnostic variables were analyzed by the Survival Estimation feature of the Cancer Data Analysis Software to obtain estimates of surviving days. Upon analysis, the number of unique treatment scenarios generated varied from 4 to 17 distinct combinations. After running each variable's analysis, the Cancer Data Analysis Control Panel displayed the mean estimated as well as actual surviving days for each unique treatment scenario based upon 10,000 to 20,000 patient samples (depending on the amount of data provided for each variable in question). The screenshot below displays the user interface of the Cancer Data Analysis Control Panel running a Comparative Analysis of Surviving Days for Treatment Factors vs. Metastases at Diagnosis at Bone.

DOI: 10.36838/v4i6.7

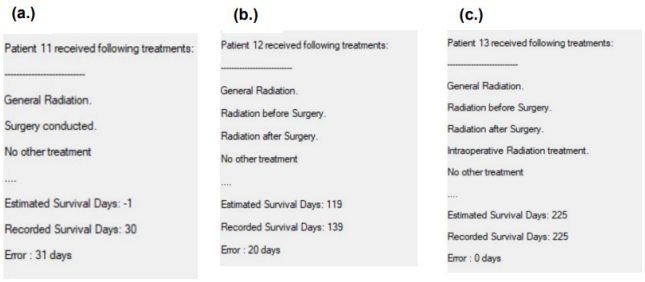


Figure 9: Three screenshots above and to the left depict the Cancer Data Analysis Control Panel's display for a Metastases at Bone Survival Estimation analysis, with (a.), (b.), and (c.) demonstrating three treatment scenarios out of 13 possible combinations for this variable with displayed estimated surviving days, actual recorded surviving days, and the error.

The Software Control Panel also displayed the resulting graph shown in Figure 10 below, popping up immediately after the text in Figure 9 was displayed. The surviving day estimation error is represented by the y-axis, whereas the unique treatment combinations are shown on the x-axis. In the case of Figure 10, there were thirteen unique combinations for the Metastases at Diagnosis at Bone variable.

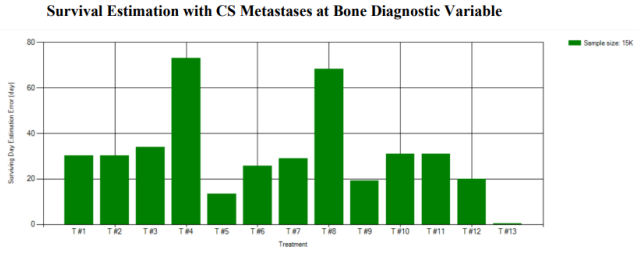


Figure 10: Accompanying the analysis in Figure 9 is this software-generated graph of treatment combinations (T1...T22) and estimated error in surviving days for CS Metastases at Diagnosis at Bone (Sample size=15,000 patients).

Table 2: Lung Cancer Survival Day Error Rates for Treatment Combination Based on Diagnostic Variable. Highlighted are the treatment combinations producing the least percent error for each variable..

Lung Cancer Survival Day Error Rates for Treatment Combination Based on Diagnostic Variable

Treatment Combination	No Diagnostic Variable	Positive Lymph Nodes	Tumor Size	Metastases in All Organs	Metastases in Lung	Metastases in Bone	Total Malignant Tumors	Primary Site of Tumor
General Radiation Only	28.62	78.97	3.21	15.73	60.34	33.95	18.17	8.04
Chemotherapy Only	30.59	24.84	14.20	24.28	23.91	9.80	44.41	10.16
General Radiation/Chemotherapy	71.25	1.73	39.31	16.55	37.37	11.85	19.75	4.60
Gen Rad/Rad After Sur/Chemo	25.83	18.57	23.50	12.65	273.53	4.64	10.06	14.03
Gen Rad/Rad Bef Sur/Chemo	15.48	2.92	*	20.92	46.60	8.76	*	11.47
Gen Rad/Rad After Surgery	31.61	63.87	*	18.88	111.11	22.32	*	*
Gen Rad/Rad Before Surgery	14.54	97.56	*	56.00	11.52	121.43	*	*
Gen Rad/RBS/RAS/Chemo	16.44	0.62	*	13.85	21.67	5.99	*	*
Gen Rad/RBS/RAS	79.26	28.89	*	8.67	41.46	14.39	*	*
Gen Rad/Intraoperative Rad	10.16	16.51	*	7.27	20.00	*	*	*
Gen Rad/RBS/RAS/IORT/Chemo	22.95	15.26	*	5.77	9.8*	*	*	*
Gen Rad/RBS/RAS/IORT	810.00	743.33	*	936.67	!	0.00	*	*
Gen Rad/IORT/Chemo	*	*	*	13.75	28.89	-	*	*
Gen Rad/Surgery/Chemo	*	*	*	4.32	70.83	21.23	*	*
Gen Rad/Surgery	*	*	*	26.67	283.33	103.33	*	*

Note: An * indicates that the treatment combination did not exist in an analysis of that variable.

Of all diagnostic variables, metastases at diagnosis in all organs provided the greatest diversity in treatment information, while Tumor Size and Total Malignant Tumors provided the least number of treatment combinations. The few treatment combinations found for the latter two indicate that there was a lack of patient data for those variables. Certain outliers were found in the data (such as the 936% error for a certain treatment combination in Metastases in all Organs) and were predominantly found in the General radiation, Intraoperative radiation, and Chemotherapy treatment combination. As such, that combination was omitted from the analysis. For this reason, median calculations were used to approximate the center of error in estimation over the mean due to outliers in data. Despite its lack of data, the Primary Site variable had the most accurate estimations with a median of 10.16% error. A close second, the Metastases at Bone variable, which had a greater quantity of data, had a median of 13.12% error. The least accurate estimations were found in the Metastases at the Lung variable, which had a median of 39.4% error. In a further statistical analysis, estimated and actual surviving days of each variable were plotted in a linear regression model to determine if a positive linear trend existed. Residual models did not conform to parametric conditions, as visible in Figure 11 (b.)

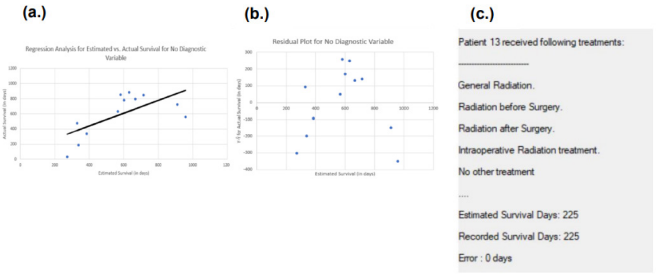


Figure 11: (a.) Least Squares Regression Model for No Diagnostic Variable shows positive trend between actual and estimated survival. (b.) Corresponding residual plot indicates non-normality of data, thus reliance on non-parametric Spearman Ranks test is needed for significance analysis.

Data Interpretation:

In terms of the Spearman Ranks Coefficient test, the strongest correlation between the ordinal variables (estimated and actual survival) was attributed to Metastases at Lung, which showed a significant degree of linear association, r_s=.88 with a p-value of .00016 for a one-tailed test (p_s> 0). Given an alpha level (α) of .05, there was very strong evidence to reject the null hypothesis (p=0). Quite strong but not as strong as for Lung, the Metastases at Bone's ordinal variables showed a significant degree of linear association, r_s=.86, p=.0002. For no diagnostic variable analyzed, there was also a significant degree of linear association, r_s=.601, p<.05, (p=.019). The Lymph Nodes diagnostic variable also had a significant degree of linear association, r_s=.59, p=.02, so the null hypothesis was rejected in this scenario as well. However, not all variables reflected significant trends between estimated and actual survival despite being positively associated with each other. The Primary Site diagnostic variable displayed a non-significant degree of linear association, r_s= .7, p=.094, and since .094>.05 the null hypothesis failed to be rejected. Similarly, the Tumor Size diagnostic variable had a non-significant degree of linear association, r_s=.8, p=.1, so the null hypothesis failed to be rejected. Finally, the greatest non-significance between ordinal variables was attributed to the number of Malignant/*In-Situ* Tumors, r_s=.2, p=.4. Thus, these results indicated that the association between the estimated and actual survival for these variables was not considered to be statistically significant for α=.05. Additionally, after median calculations of estimated and actual surviving days for each diagnostic variable, it was found that Metastases at Bone had the least median estimated and actual survival (201 and 220 days respectively). This is

to a 2017 Danish population-based cohort study conducted on patients with various primary cancers, which found that one-year survival after bone metastasis diagnosis was lowest in patients with lung cancer.[22] Moreover, the fact that Metastases in Lung had the past accurate estimations in Table 2 is understood in the context that the lung is the origin of lung cancer, so the patient likely had existing lung metastases serving as a confounding variable in survival estimation.

Discussion

Based on the results from the Spearman Ranks Test measuring the reliability of the association between estimated and actual surviving days, the original hypotheses were supported by the control, as well as the Lymph Nodes, Metastases at Bone, and Metastases at Lung variables. This indicates that the relationship between actual and estimated survival days was significant and likely to persist in future trials when considering these factors. Thus, it can be concluded that the lowest (highlighted) error rates shown in Table 1 for each of the variables will most likely repeat if additional lung cancer patient samples were added. For example, the control group's passing of the Spearman Ranks test indicates that the lowest error in surviving days matching the best combination of General Radiation and Intraoperative Radiation in Table 1 will persist in new samples. As such, this specific set of treatments should be implemented in the place of general "combinations of" chemotherapy, radiation, and surgery described by the CDC as current methods to combat lung cancer. As for Lymph Nodes, treatments have traditionally involved radiation to the chest, as well as chemoradiation for patients with limited-stage small-cell lung cancer. In line with the current therapeutic approach, chemotherapy should be paired with general radiation and intraoperative radiation conducted before and after surgery to yield the most accurate surviving days. For Metastases at Bone, which also passed the Spearman Ranks Test, the combination of all radiation-related treatments will most likely lead to more accurate estimations of survival by the software in future trials. Thus, multiple forms of radiation should replace current methods of surgery, radiotherapy, and bisphosphonates used to treat bone metastases. Similarly, for Metastases in Lung, all radiation-related variables and chemotherapy were included in the best treatment combination. As such, they should be used in conjugation to improve patient survival. In current literature, the best treatments for lung metastases include chemotherapy, immunotherapy, or radiation therapy, with "a combination of these" ideal according to the American Cancer Society. However, results illustrate that the combination of chemotherapy and general radiation with more sophisticated forms of radiation like intraoperative radiation at specific times (before and after surgery) will encompass a more holistic approach to therapy. The significant inferential test result obtained by the Metastases at the Lung variable is also understood in the context of lung metastases occurring at the site of cancer. Hence, the relationship between actual and estimated survival for lung cancer considering this variable will be more reliable.

However, since the Primary Site, Tumor Size, and Malignant Tumor variables did not pass the Spearman Ranks test, they are not reliable indicators of accurate survival estimation. However, there were key limitations that could have reduced the accuracy of the results. Most directly related to the non-significance of these variables is the lack of patient data. A prominent pattern among these three variables is the lack of unique treatment combinations in Table 1. As such, it is understood that a lack of unique patient data covering all treatment combinations resulted in the non-significant Spearman results for these variables. A key limitation impacting all diagnostic variables, however, is the non-linearity of the data. Due to the non-linear nature of cancer prognosis and confounders, survival estimation cannot be perfectly linearly associated with patient variable data. Although matrices were normalized as part of the procedure to "linearize" the data through retrieval of mean surviving days for patients with similar treatment conditions, this process was not able to completely linearize the data due to confounding variables unaccounted for by the SEER data. Previous studies that have estimated cancer survival rates in conjunction with patient variables such as those studies by respective authors She *et al.*, Lai *et al.*, and Burki *et al.* have utilized machine learning or deep learning algorithms to account for this limitation. However, due to the lack of patient data, Machine Learning was not feasible for this study. Finally, inherent biases in SEER data collection were implied through the signing of the SEER Data Use Agreement form.

Thus, the original hypotheses about the single-variable, differential analysis, and survival estimation portions of the study were partially true. Reflecting on the single-variable analysis prediction, it was true that the number of surviving days of a respiratory cancer patient was determined by the various social, genetical, and diagnostic variables- evident by the variety of means and medians in the two box plot examples provided in Figure 7 for the Lymph Nodes and Metastases at Lung variables. However, in terms of the differential analysis, the mathematical model did not reflect a linear system despite normalization due to the presence of negative treatment and diagnostic coefficients as seen in Table 1, as well as the reasons provided in the previous paragraph. As for survival estimation, the percent change between estimated and actual survival was not quite zero according to Table 2. However, in the case of the least error combination for the Lymph Nodes variables, the error came close at 0.62%.

Conclusion

In synopsis, the Cancer Data Analysis Software developed as part of this study serves as a valuable tool to healthcare professionals for the prognosis and assignment of lung cancer treatment by issuing a personalized, least-error treatment plan for patients after considering unique diagnostic variables and combined treatment options. Stemming from the significant relationships between actual and estimated survival for certain diagnostic variables supported by the Spearman Ranks and p-score significance tests, patients exhibiting these variables can be assured that the least error treatment combination will provide the most accurate surviving days- thus prolonging survival. A "personalized outlook" was established through consideration of unique diagnostic and treatment variables

DOI: 10.36838/v4i6.7

of existing SEER patients, including but not limited to diagnosis-related variables such as Lymph nodes, metastases, and tumor size, and specific treatments like Radiation Before Surgery. However, it is acknowledged that outlook cannot be completely personalized since survival estimation only focused on diagnostic and treatment-related variables, excluding the social/genetic variables originally used in the single-variable analysis. Nonetheless, through consideration of complex treatment combinations representing the true situation of many lung cancer patients, estimates are personalized in that they are not limited to analysis of a single therapy (not a likely scenario for many patients). This software, paired with the TNM system and existing screening processes, can give doctors the leverage needed to provide a well-rounded approach to prognosis through statistical analysis-driven validation. Moreover, the quality of life of lung cancer patients will be improved as patients are no longer dependent on doctors' potentially subjective prognoses fueling the risk of overdiagnosis. Specifically, by knowing which treatment combination best prolongs survival, patients can be informed of which costly treatments to avoid in the light of "overtreatment" concerns discussed by Dr. Martin Makary, a professor of Surgery at Johns Hopkins University School of Medicine in Baltimore.[23] This study holds especially great value for patients with early-stage lung cancer as a study conducted by the Fox Chase Cancer Research team at Temple Health, Philadelphia found that individuals with early-stage lung cancer and small tumors were typically overtreated.[24] Finally, the researcher hopes that the Cancer Data Analysis Software Application will become an open-source resource for other cancer researchers to utilize for their analyses, allowing for the choice of a particular diagnostic variable and relevant analysis type (single variable analysis, differential analysis, or survival estimation).

Future Study:

To improve this year's research, incorporating additional variable data such as genetic data in the form of molecular subtype or health condition, lung cancer type (non-small cell vs. small cell), and stage of the tumor could provide a more holistic view of most effective treatment options while reducing the presence of confounding variables. Additionally, through gathering a greater sample pool for variables limited in data such as Primary Site and Total Malignant/*In-Situ* Tumors, the number and types of treatment combinations across all diagnostic variables can be made homogenous. This would allow Spearman Correlation tests to not only be conducted for each diagnostic variable type, but also each treatment combination, leading to a complex understanding of which variable and treatment combinations respectively would be most impactful on survival. Moreover, through a more effective linearization method, the software will be able to make more accurate predictions. In the future, the building of new software modules within the Cancer Data Analysis Software for the analysis of Colon cancer and Pancreatic cancer, which correspond to 22% and 9% of total cancer deaths in 2021 respectively, could allow for more statistical analyses not only within but between cancer types, contributing to a more complex understanding of cancer survival prognosis.

Acknowledgements

Melissa Erdem was the sole author of this research paper. She was first interested in oncological research in her sophomore year of high school when she became aware of a family friend's struggle with a respiratory cancer diagnosis.

References

1. National Cancer Institute. Common Cancer Sites - Cancer Stat Facts. https://seer.cancer.gov/statfacts/html/common.html.
2. National Cancer Institute. Cancer of the Lung and Bronchus - Cancer Stat Facts. https://seer.cancer.gov/statfacts/html/lungb.html.
3. Centers for Disease Control and Prevention. How Is Lung Cancer Diagnosed and Treated? https://www.cdc.gov/cancer/lung/basic_info/diagnosis_treatment.htm.
4. National Cancer Institute. Staging. https://www.cancer.gov/about-cancer/diagnosis-staging/staging.
5. The Radiology Assistant: TNM classification 8th edition. https://radiologyassistant.nl/chest/lung-cancer/tnm-classification-8th-edition.
6. International Association for the Study of Lung Cancer. Highlights of the 8th Edition of the TNM Staging System: Practicalities and Tools https://www.iaslc.org/iaslc-news/ilcn/highlights-8th-edition-tnm-staging-system-practicalities-and-tools (accessed Sep 26, 2021).
7. Yun, J. K.; Lee, G. D.; Kim, H. R.; Kim, Y.-H.; Kim, D. K.; Park, S.-I.; Choi, S. Validation of the 8th Edition of the TNM Staging System in 3,950 Patients with Surgically Resected Non-Small Cell Lung Cancer. Journal of Thoracic *Disease* **2019**, 11 (7), 2955–2964. https://doi.org/10.21037/jtd.2019.07.43.
8. Hattori, A.; Takamochi, K.; Oh, S.; Suzuki, K. New Revisions and Current Issues in the Eighth Edition of the TNM Classification for Non-Small Cell Lung Cancer. *Japanese Journal of Clinical Oncology* **2018**, 49 (1), 3–11. https://doi.org/10.1093/jjco/hyy142.
9. Zurrida, S.; Veronesi, U. A New TNM Classification for Breast Cancer to Meet the Demands of the Present and the Challenges of the Future. *Women's Health* **2011**, 7 (1), 41–49. https://doi.org/10.2217/whe.10.82.
10. Ge, Z.; Heitjan, D. F.; Gerber, D. E.; Xuan, L.; Pruitt, S. L. Estimating Lead-Time Bias in Lung Cancer Diagnosis of Patients with Previous Cancers. *Statistics in Medicine* **2018**, 37 (16), 2516–2529. https://doi.org/10.1002/sim.7691.
11. Gill, R. R.; Jaklitsch, M. T.; Jacobson, F. L. Controversies in Lung Cancer Screening. *Journal of the American College of Radiology* **2016**, 13 (2), R2–R7. https://doi.org/10.1016/j.jacr.2015.12.005.
12. Wegwarth, O.; Gigerenzer, G. Overdiagnosis and Overtreatment. *JAMA Internal Medicine* **2013**, 173 (22), 2086. https://doi.org/10.1001/jamainternmed.2013.10363.
13. Lambden, J.; Zhang, B.; Friedlander, R.; Prigerson, H. G. Accuracy of Oncologists' Life-Expectancy Estimates Recalled by Their Advanced Cancer Patients: Correlates and Outcomes. *Journal of Palliative Medicine* **2016**, 19 (12), 1296–1303. https://doi.org/10.1089/jpm.2016.0121.
14. Lynch, C. M.; Abdollahi, B.; Fuqua, J. D.; de Carlo, A. R.; Bartholomai, J. A.; Balgemann, R. N.; van Berkel, V. H.; Frieboes, H. B. Prediction of Lung Cancer Patient Survival via Supervised Machine Learning Classification Techniques. *International Journal of Medical Informatics* **2017**, 108, 1–8. https://doi.org/10.1016/j.ijmedinf.2017.09.013.
15. She, Y.; Jin, Z.; Wu, J.; Deng, J.; Zhang, L.; Su, H.; Jiang, G.; Liu, H.; Xie, D.; Cao, N.; Ren, Y.; Chen, C. Development and Validation of a Deep Learning Model for Non–Small Cell Lung Cancer Survival. *JAMA Network Open* **2020**, 3 (6), e205842–e205842.
16. Doppalapudi, S.; Qiu, R. G.; Badr, Y. Lung Cancer Survival Peri

od Prediction and Understanding: Deep Learning Approaches. *International Journal of Medical Informatics* **2021**, 148, 104371. https://doi.org/10.1016/j.ijmedinf.2020.104371.
17. Lai, Y.-H.; Chen, W.-N.; Hsu, T.-C.; Lin, C.; Tsao, Y.; Wu, S. Overall Survival Prediction of Non-Small Cell Lung Cancer by Integrating Microarray and Clinical Data with Deep Learning. Scientific Reports 2020, 10, 4679. https://doi.org/10.1038/s41598020-61588-w.
18. Burki, T. K. Predicting Lung Cancer Prognosis Using Machine Learning. The Lancet Oncology 2016, 17 (10), e421. https://doi.org/10.1016/s1470-2045(16)30436-3.
19. SEER Incidence Database - SEER Data & Software https://seer.cancer.gov/data/#:~:text=SEER%20collects%20cancer%20incidence%20data%20from%20population-based%20cancer.
20. National Cancer Database (NCDB) | sgim.org https://www.sgim.org/communities/research/dataset-compendium/national-cancer-database-ncdb#.
21. Duggan, M. A.; Anderson, W. F.; Altekruse, S.; Penberthy, L.; Sherman, M. E. The Surveillance, Epidemiology and End Results (SEER) Program and Pathology: Towards Strengthening the Critical Relationship. *The American journal of surgical pathology* **2016**, 40 (12), e94–e102. https://doi.org/10.1097/PAS.0000000000000749.
22. Svensson, E.; Christiansen, C. F.; Ulrichsen, S. P.; Rørth, M. R.; Sørensen, H. T. Survival after Bone Metastasis by Primary Cancer Type: A Danish Population-Based Cohort Study. *BMJ open* **2017**, 7 (9), e016022. https://doi.org/10.1136/bmjopen-2017-016022.
23. Szabo, L. So Much Care It Hurts: Unneeded Scans, Therapy, Surgery Only Add To Patients' Ills https://khn.org/news/so-much-care-it-hurts-unneeded-scans-therapy-surgery-only-add-to-patients-ills/
24. Patients with Clinically Diagnosed Early-Stage Lung Cancer Potentially Overtreated https://www.foxchase.org/news/2016-04-06-Clinically-Diagnosed-Early-Stage-Lung-Cancer-Potentially-Overtreated.

Author

Melissa Erdem is a high school senior and intends to major in Microbiology and Cell Science on a Pre-Medical track at the University of Florida. Apart from conducting biomedical research, Melissa is also an avid violinist and intends to join UF's Symphony Orchestra. She also enjoys spending time with her sisters, traveling, and assisting non-profit organizations.

DOI: 10.36838/v4i6.7

IJHSR
ijhighschoolresearch.org
■ REVIEW ARTICLE

The Diversity within Neurodiversity: Recommendations for a New Approach to ASD Treatment Priorities

Isabelle Hsu
Irvington High School, 41800 Blacow Road, Fremont, CA, 94538, USA; isabellehsu123@gmail.com

ABSTRACT: This paper discusses the complicated and intertwined factors and controversies surrounding neurodiversity, with a particular focus on Autism Spectrum Disorder (ASD). Not only is the nature of neurodiversity multifaceted and sometimes hard to characterize, but there are also multiple factors and controversies affecting treatment planning. Due to the complex nature of ASD, it is proposed that the standard practice of basing treatment priorities heavily on parents' major concerns should be reconsidered and that a more systematic and objective education program and categorization approach should be developed in future research.

KEYWORDS: Behavioral and Social Sciences; Clinical and Developmental Psychology; Neurodiversity; Autism Spectrum Disorder (ASD); ASD Treatment Priorities.

■ Introduction

Neurodiversity has become a hot topic in recent years and has stirred up a great number of controversies. The term "neurodiversity" was first introduced by Judy Singer in 1998 as part of the autism rights movement.[1] It implies that neurological differences and the resulting behavior should be treated and respected as forms of natural human variation.[2] In other words, although some conditions might appear to be deficits and "inconvenient" to have because they make it harder for individuals to fit into the norms of society, these conditions should not be treated as "abnormalities" or illnesses that need to be cured. Researchers have further found various strengths, or natural talents, associated with having these neurological conditions. Specifically for individuals with Autism Spectrum Disorder (ASD), which contributes to a large portion of the neurodiverse population and seems to attract the most attention in research, talents in music, math, memory, and spatial skills are among the many strengths that have been commonly found.[3] However, many ASD individuals are unable to utilize their special strengths due to a variety of reasons, and it is certainly a loss not only to them but also to our society as a whole. As a result, this paper not only explains the less-apparent sides of ASD, such as some of the common strengths that tend to come with it and controversial topics surrounding it but also proposes the development of education and communication tools to help patients and families make well-informed decisions together with their clinicians.

■ Discussion

I. Complicated Nature of "Neurodiversity":

Neorodiversity is a very complicated topic from many aspects.

A. A Diversity of Models

There are multiple models that view the nature of neurodiverse conditions from very different perspectives. The most prominent framework for understanding autism, historically, is a medical model where ASD is treated as a disability. While the medical model takes a more traditional view to see ASD as deficits to be treated and corrected, a strength-based model focuses on developing such individuals' areas of strength. It considers a more comprehensive panel of criteria, such as talent in music or even other daily tasks the individuals are better at, that could help them excel and have a sense of accomplishment. In turn, these strengths can be used to motivate them to work harder on other necessary skills and aspects.[4] In addition, there is a social model of disability that sees the challenges of these individuals as the responsibility of society as a whole. In other words, society fails to see these variations in brain functions as normal in humans and does not provide enough flexibility and space for these variations.[5] Whether it is the medical, social, or strength-based model, they are each with a very different perspective, and the treatment decisions based on the different models will likely be very different.

B. Identity-First vs Person-First Language

Besides the different models mentioned above, there are also controversies in terms of how to refer to these individuals. While self-advocates usually prefer being addressed in identity-first language ("autistic individual"), the person-first ways, such as "individuals with autism," are usually preferred by advocates for them such as parents.[2] One of the reasons that many self-advocates prefer identity-first language is because they feel this identity should be something to be celebrated. One analogy that is often heard among those in favor of identity-first language is, as a proud American, people would say that they are American citizens, instead of people with American citizenship, and therefore, instead of individuals with autism, they would like to identify themselves as autistic individuals. For the purpose of this paper, person-first and identity-first language will be used alternatively to represent views from both groups.

C. "High" vs "Low" Functioning Autism

Furthermore, there is also controversy regarding the labels of "high functioning" and "low functioning" autism. Usually, so-called high functioning refers to people who are not cognitively compromised and have a normal Intelligence Quotient

DOI: 10.36838/v4i6.8

(IQ) range of 70 or above.[2] This group usually exhibits behavior that is closer to the "normal," or the neurotypical, majority. The other group, the low-functioning ones, has an IQ range of 70 or lower. However, questions arise about whether IQ testing is even an appropriate measure to determine intelligence levels for ASD individuals. For example, for non-verbal, occasionally also referred to as non-speaking to emphasize impairment over being non-functional, individuals on the spectrum, sometimes even the verbal ones, such testing designed for the neurotypical majority could end up being a very biased way of measuring their cognitive ability.6 In other words, non-verbal ASD individuals can score very low on IQ testing due to their inability to communicate, not actually due to their low intelligence levels. As a result, giving the high and low functioning labels is not only controversial but could be a very poor indicator of intelligence in the autistic population

D. Asperger's Syndrome in ASD Diagnosis

Moreover, Asperger's Syndrome, which used to be a separate diagnosis from ASD, has now been part of the spectrum since the release of the DSM-5 in 2013.[2] Individuals with Asperger's, although experiencing various challenges due to the neurological differences, are usually said to be "high-functioning," and sometimes even "twice-exceptional" with IQ measured over the levels 120 or even much higher. From the perspective of the medical model, although Asperger's' patients experience the same social and communication deficits, some argue that including it in ASD misrepresents the condition because the level of "deficits" or "disability" associated with Asperger's is usually much lower than others in the spectrum.[2] Some go further saying such misrepresentation jeopardizes other ASD patients' chance of getting disability benefits or accommodations because it gives the general public the wrong image that all individuals on the spectrum are savants since some individuals with Asperger's are especially talented in certain ways, such as memory and math.[3] An example that is often referred to is Barry Levinston's 1988 film, Rain Man, and for the general public who are less experienced with this group of the population, this might become the only source of information and impression they have for people on the spectrum. Whether or not Asperger's should be included in the spectrum diagnosis is beyond the scope of this paper, but this controversy further highlights the complex nature of ASD.

E. Altered Perceptions Over Time

It is also interesting to note that while previous research has estimated that 70-80% of ASD individuals were cognitively impaired with IQ levels below 70,[7,8] more recent 2014 CDC studies indicate only 31% are estimated to have such impairments.[9] One reason for the discrepancy could be due to the inclusion of Asperger's in the ASD diagnosis, which increases the number of normal to high IQ patients. Another reason may be that with more research and knowledge in the field, many of the originally "less obvious" cases with less noticeable symptoms are now being correctly diagnosed. Now that individuals with fewer cognitive impairments are being correctly diagnosed with ASD, the prevalence of impairment in the population has reduced. This trend certainly helps reduce the overall stigmatization of ASD, where now a significantly lower portion of diagnosed patients are being labeled as cognitively impaired. On the other hand, it again brings up concerns about its impacts on patients' chance of obtaining disability benefits, because ASD could then be perceived by the general public as just a relatively minor variation in terms of communication and social skills

II. Factors and Considerations in Treatment Planning:

As can be seen, there are many controversies around the ASD condition. However, even more arise regarding the best practices of assessment and intervention for these patients.

A. Intervention Needs Dependent on the Nature of the Behavior

First, it is worthwhile noticing that even strength-based and social-model advocates would agree that certain ASD behaviors undoubtedly need interventions.[10] The most obvious cases are individuals with self-inflicted injuries or physical aggression towards others. For the cognitively affected and even some of those with more typical cognitive abilities, it is usually the consensus to emphasize the importance for them to develop the necessary life skills to help themselves with self-care and to achieve as close to independence as possible.[8,11] However, in terms of certain social behavior, it is now divided whether certain behavior previously considered "socially inappropriate" should now be taken as acceptable through more awareness in the general public.[12] The most obvious example is stimming, which often involves repetitive physical movements or sounds, for ASD individuals. Individuals on the spectrum stim for purposes such as releasing sensory overload so they can calm themselves down and perform better.[13] While such behavior does not harm anyone, it is often perceived as simply weird or "inappropriate" because it is not done by others. Then questions arise as to whether such behavior needs to be corrected for autistic individuals to be more easily accepted in society.[10,13] Because correcting such harmless behavior is counter-productive to sensory overload, researchers have found it unnecessary and even harmful to try to correct it.[13]

B. Extent of Being Beneficial and Practical in Inclusion

Nevertheless, whether it is practical not to treat some relatively non-harmful behaviors is sometimes not a simple question. ASD itself is complicated, not to mention that it also often comes with other conditions such as ADHD, OCD, and Tourette's.[14,15] Different conditions indeed bring different possible strengths to the individual, and the goal of neurodiversity advocacy is often for society to have enough awareness to accommodate their various special needs so that these individuals will get a chance to develop and make use of their strengths.[12] However, today's society is not as accommodating as it ideally should be, and therefore, whether ASD individuals should "fix" some of their neurodiverse behavior so that they can get a better chance to fit in and use their strengths in this society has become a question of debate.

For example, while it would be ideal for the individual to get a customized 1-on-1 education to develop those strengths, it is often not the most budget-friendly. Including students with special needs in regular classrooms, or called "inclusive

DOI: 10.36838/v4i6.8

classrooms," is the more common approach to developing higher-level skills because these skills, especially academic ones, are taught in those classrooms. While teachers in inclusive classrooms can be trained to use tools and strategies such as assistive technology, seating arrangements, buddy systems, and modified lesson plans, to make the class more inclusive, children with autism might still have a hard time paying attention and learning in such a classroom setting, and some autistic behavior, such as difficulties taking turns and tendencies to speak loudly or off-topic, might make it impractical for them to be included in such a classroom as they may "interfere with instruction or attempts at inclusion."[16] Also, even if under a preferred case where a teacher is trained to handle students with various conditions in an inclusive environment, a limited budget might not always allow the class size to be small enough for such a teacher to implement all the necessary accommodations while taking care of a large classroom of students. Sometimes it works when the budget allows a trained aide to be with the special-education student in the regular classroom. However, this may or may not be enough because such a student may require a structured approach with "explicit teaching" and an "experienced, interdisciplinary team of providers."[14]

Therefore, it is often not practical to expect that students with special needs can completely "be themselves" while the teacher can also keep a sizable classroom of students in a proper environment for learning without frequent disruptions. Under such circumstances, a dilemma arises as to whether students with such conditions should learn to fit into the norm so that they can be "included" in a regular/inclusive classroom setting to learn more and to develop their strengths, or if they should be allowed to be themselves and be separated in special-education classrooms that often have lower academic expectations.

C. The Two Sides of Masking Behavior

Another relevant issue is masking, which is when ASD individuals copy the behavior of neurotypicals to make themselves appear to be more "normal." It is not unusual for ASD individuals to learn to mask, or camouflage, their conditions, either knowingly or unknowingly to themselves.[17] Some might be aware of their atypical thoughts and behavior and intentionally learn to hide them by pretending to behave like the neurotypical. Others naturally learn to copy others' behaviors to act to the norm, even though such actions might not make sense to themselves or sometimes are even so unnatural to the point that it becomes psychologically painful in the long term.[17] Masking behavior does make autistic individuals appear to be more normal and often helps them fit in socially. However, these experiences can be "exhausting, isolating, damaging for their mental and physical health, identity, and acceptance of self, creating unreal perceptions and expectations of their abilities for others, and in some cases led to a delay in formal diagnosis or a mental health crisis."[17]

Although due to the long-term stress and the various negative psychological impacts of "pretending," masking is not recommended by many professionals, the group of ASD individuals who learn to mask tend to possess higher cognitive ability and feel they have more involvement or "access" to the social world.[17] Without such higher cognitive ability, it would not have been possible for them to figure out the difference and to copy the atypical behavior and responses in the first place. Also, they tend to be the group who will have a higher chance to live independently, feel a sense of achievement, and have relatively "normal" or even "successful" lives, whether it is due to their higher cognitive ability or as a result of their success from looking "normal" to fit in.[17] Ideally, society should learn to accept a diverse range of behavior due to neurological differences, but in an imperfect world, it is sometimes still a dilemma whether they should be encouraged to mask their conditions at the cost of long-term mental well-being.

D. Issues with Diagnosing "Twice Exceptionality"

There is also a group of especially high IQ ASD individuals who are often called "twice-exceptional," people who are special both in terms of "deficits" and talents. Depending on how visible the autistic traits are, they tend to go undiagnosed for ASD and unidentified for being gifted.[18,19] First, their exceptionally high IQ, typically greater than 130, often makes their challenges less visible.[18] One might think that this is not a big deal because the person acts just like the rest of the ones around, but in reality, it is often frustrating. For example, such an individual might constantly feel that they are talented or very capable of outsmarting their peers but get frustrated at the fact that they are never able to demonstrate that in reality because their talents are "crippled" or limited by their ASD challenges such as social or communication skills that make those advantages impossible to be seen or identified. On the other hand, not being identified for their ASD challenges because of their especially high cognitive abilities also places such individuals in constant confusion as to why the world around them seems to operate in ways that do not make sense to them. As mentioned above, although individuals with ASD might not understand certain things that neurotypical individuals understand and do naturally, some of them learn to mimic the behavior of neurotypicals well enough that they end up undiagnosed. There are many such cases when ASD individuals are finally diagnosed with ASD later in life, they experience tremendous relief because they can finally find out why they have been feeling that way all their lives.[1] As a result, while such individuals might appear to function adequately alongside their neurotypical peers, it is important that both the ASD and the giftedness are identified. Moreover, as a note, currently twice exceptionality is based on higher IQs, but there are also ASD individuals who either possess natural strengths in areas not measurable by IQs or are limited by their symptoms such as their verbal abilities that make them unable to demonstrate such high IQ levels, further complicating the assessment of twice exceptionality and their subsequent treatment and development plans.[6]

E. Dilemmas in Allocating Resources

Although many ASD individuals possess natural talents in certain areas, dividing resources such as time and budget between working on ASD symptoms and developing the strengths is often a real challenge. For example, it might be very important for speech therapy to take place regularly for a non-verbal ASD individual, but it will also be a waste if such

an individual's time is mostly devoted to therapy and not to his exceptional music or math talent that could potentially bring much higher self-esteem and even a successful career. On the other hand, if time and resources are devoted to developing music or math talent, the individual might lose the opportunity to spend enough time to be able to communicate better to live a relatively independent life.

F. Facing Multiple Conditions

Having multiple conditions, whether directly related to ASD in nature or other unrelated physical or psychological conditions, also impacts treatment planning. Besides having multiple neurological conditions and/or being cognitively compromised, other physical or psychological challenges are also commonly found among ASD individuals. As mentioned in section II.B, some ASD individuals have other neurological conditions such as ADHD and Tourette's Syndrome and/or medical challenges such as other chromosome abnormalities.[14] Studies have also shown that they also tend to develop other psychological disorders such as depression and anxiety due to the constant stress, either from having to mask their behavior to fit in or from never being able to feel accepted or having a sense of belonging in community.[20, 21] In addition, these individuals also tend to have more medical conditions related to gastrointestinal issues, sleep disorders, and seizures.[22] These additional conditions certainly make intervention decisions a more complicated process.

G. Environmental Factors

Environmental factors such as resources available to the individual can play an important role in interventions as well.[23] Family support, financial stability, state or public education funding, as well as available specialized physicians nearby also contribute to what and how interventions will take place. For immigrant families, finding specialists and therapists who can speak their language and communicate in a more culturally sensitive context will be important. Family status, such as domestic conflicts, as well as support from extended families and grandparents also play a factor in determining the frequency and the extent of the intervention. Moreover, if a parent or a sibling also has special needs, or if there are new additions to a family, it will also affect how much focus could be placed on the particular ASD individual. Furthermore, the particular county's funding for intervention programs, as well as the school district's commitment to special education, all play vital roles in interventions. As a note, since COVID-19, many services are available online through telehealth platforms such as Zoom, making many previously inaccessible services and therapists available to patients not in the same geographic area.[24] Although it is unclear whether the use of telehealth platforms will continue to be a trend after the pandemic, it has certainly opened the doors for seeking professional services that were physically out of reach.

H. Changed Priorities Over Time

Finally, it is critical to assess treatment progress regularly and use that as a factor in future treatment planning. Sometimes progress such as the ability to retain previously learned skills should be used to determine which types of interventions are appropriate. Reaching a certain goal might not mean a shift of focus should be implemented. The same treatment plan and goal could still be considered of higher priority especially if it involves great effort and time to achieve the level of success. An example could be academic performance in a certain subject. Reaching a certain grade level in a particular subject might not mean the effort should stop there. On the contrary, consistent effort should be made to ensure success in the long run, and the positive experience will likely foster higher self-esteem and motivate the individual to make more effort either in the same or other areas. As another example, certain unwanted behavior such as aggression toward others might be regarded as an important milestone in intervention, but even if it has been stopped earlier in the treatment, such concepts should continue to be reinforced during the intervention as prevention.

As can be seen, there are a variety of factors unique to each patient that could affect treatment planning and prioritization. Although some factors might play a more significant role in one patient's situation over another's, not having a comprehensive view when making the planning decisions could very possibly jeopardize the best interest of the patient and the overall outcome.

III. Controversies in Treatment Methods:

Early diagnosis and intervention are said to be critical to ASD patients.[25] However, sometimes symptoms, especially in the less severe and thus less obvious cases, can be overlooked before a child is sent to preschool, and special education programs might not be available until the child is in a public school system. Once a child is identified as in the spectrum, a variety of interventions can be done. They may include play-based therapies to learn things such as recognizing facial expressions, speech therapy, group sessions to develop social skills, and some might involve medications as well as physical therapy for related motor skill development. However, there are also controversies surrounding the treatment approaches themselves.

Some of the traditionally and commonly used therapies are being reconsidered. For example, while Applied Behavioral Analysis (ABA) therapy has been very popular for correcting autistic behavior over the past decades, some researchers and advocates criticize the way the treatment is done as abusive and oftentimes similar to treating humans like animals.[26] Some ABA therapy does involve forms of punishment as negative reinforcement to produce desired behavior, and these opponents argue that humans should not be forced into producing behavior to suit society's expectations. There is also a study that shows that ABA participants often develop Post Traumatic Stress Disorder well into their adulthood.[27] However, depending on how such therapies are conducted, there are also ABA therapists that use strictly positive reinforcement to produce desired results. Therefore, it is probably not the name of the therapy, but exactly how it is done under different circumstances that matters.

In addition, although ASD is found to have strong genetic links because they tend to run in families, the real causes are still quite unclear.[8,14] There are always forms of alternative treatment plans that some feel helpful but are never medically proven to be effective. Some of those alternative treatments could simply be unhelpful for most people (such as a gluten-free diet), but others could be potentially harmful to the

DOI: 10.36838/v4i6.8

"higher functioning" ASD cases.[14,15] Because most symptoms of autism cannot be changed with psychotherapy, it is an often-overlooked component of the treatment plan.[15] Parents are also likely to regard psychotherapy as something unnecessary because it does not usually solve any of the core symptoms of autism or eliminate any visible "problems."[15] However, as discussed before, many high-functioning autistic individuals are more prone to mental health issues such as depression and anxiety.[21,32] As a result, it is often in the best interest of the patients to have such psychotherapy, even though it does not make their autism symptoms "better."[21,32]

C. When Parents' Interests Need to be Prioritized

Questions arise whether or not a parents' interest should ever be considered before the child's best interest. In general, the autistic individual's best interest should always be the priority. However, there are certainly circumstances under which the parents' best interest should be placed higher. For example, studies have shown that parents are often at "increased risk for depression or stress-related illness as a result of the unique problems inherent in living with a child with a serious disability."[14,33] One might argue that parents' mental health issues should be separated from the child's issues, instead mixing the parents' needs with the child's treatment needs. In a perfect world, that should probably be the case. However, sometimes it would be more practical to target a particular behavior that severely impacts parents' well-being, especially when a parent is unable or unwilling to seek additional medical help for themselves in the midst of the child's demanding treatment needs. After all, because parent involvement is oftentimes critical in the treatment process, the well-being of the parent directly makes a difference in the effectiveness of the treatment plans. Therefore, sometimes it might make sense to consider parent-preferred "behavioral improvement" a priority, at least temporarily, as it could possibly reduce stress and depression symptoms for the parents, making it possible for parents to be available and helpful during the course of the treatment.

V. Future Research: A More Systematic and Objective Approach for Setting Up Intervention Priorities:

When planning treatment for autistic patients, even for some autistic adults, parents' preference often plays an important role in determining what interventions will be conducted and the details like how and when they will be conducted. It is for obvious reasons, such as parents being the ones that spend the most time with these ASD patients, and since these patients might not be able to express their preferences, parents naturally become their spokesperson. In addition, parents are also the transportation providers for appointments and therapies, not to mention they often also need to continue to reinforce what has been learned during therapy sessions at home. Therefore, their opinions on intervention are often highly valued in the decision-making process. However, this paper challenges this intuitive approach which relies heavily on parents' preferences, and specifically on the assumption that parents are equipped to make the best choice for their autistic child.

As discussed, ASD is an overwhelmingly complex issue from many perspectives, so even though parents seem to be the ones that know their child best, they might not see the whole picture and have comprehensive knowledge of the field when they make intervention decisions for their child. As a result, this paper proposes the use of a more systematic and objective approach with parent education (patient education as well, if the patient is old enough and cognitively able to understand the content) as the first steps to generate intervention and treatment priorities. Future research should aim for developing parent education, in a format and language that is designed to be easily understandable to non-professionals, as well as tools designed to help draw out treatment priorities based on objective criteria, so that treatment planning will no longer be based on parents' intuition or major complaints. Of course, once the parents are presented with the result and explanations of such objective assessment that relies on research-based criteria instead of parents' intuition, communication should be done between the clinician, the parents, and if possible, the patients, so that then their informed preference will still be taken into consideration for setting up the actual treatment plan. A more comprehensive and objective approach that still involves the parents and the patients but equips them with enough knowledge to make informed decisions, will likely help pinpoint the most critical areas to work on, making the treatments more likely to benefit the patients.

VI. Recommendations for Improved Communication and Decision Making in Treatment Planning:

This paper would like to propose two specific ideas for future research that could help simplify the process of communicating such complicated issues surrounding ASD to parents and patients without much background knowledge.

A. Education Program for Parents and Patientss

An education program should be developed to explain the controversies and complexity of the issue surrounding ASD to parents and patients. There should be a basic and an elaborated version of the program, with the shorter version ranging from 2-4 hours that cover the most important issues, such as the basics of neurodiversity and strength-based approach, balancing between tackling parents' major complaints vs issues that benefit the patients directly, parents and patients' psychological health, and commonly discussed questions at different life stages. Such a program aims to give parents and patients a science-based quick overview of what is important but might not be very intuitive to them, for example, when focusing on what bothers them the most might not benefit the patient the most. Also, a longer program that aims to provide more knowledge and resources on a continuous basis for parents and patients should be made available, especially to those who are able to invest time to better equip themselves with higher levels of knowledge to make more informed decisions in their future treatment plan modifications. This could also have a ripple effect in the community, as some parents may have the time to pursue more in-depth education and can share advice and resources with other families. Given the complexity of autism and recent advances in research, there is so much potential to develop resources based on the data that exist and the importance of passing this knowledge to

the decision makers (e.g., parents) in a structured format that is easy to understand.

Depending on available funding and practicality, such a program can be potentially made into quite a few different formats. It could be a series of sessions offered by clinicians to their own patients' parents, community workshops organized by practitioners in the same geographical area with or without a live discussion and support component, or even Zoom classes hosted by trained professionals across a large geographical area that can be made readily available to parents of newly diagnosed ASD patients. If resources are limited, a program as simple as a link to a series of online videos provided to the parents could also work, although it might not be as effective as having direct guidance from a live professional who would be able to address questions. Bear in mind, however, that delivering such education to the parents in a timely manner before determining a treatment plan should be an important goal of the program.

B. Detailed Categorization of Patients' Profile

Since the term "on the spectrum" is such a broad term that covers a huge variety of complicated symptoms, not to mention other commonly seen issues surrounding it, it will be helpful for some kind of a category code system to be developed to help parents and patients get a more concrete sense of their overall situation. First, a chart with letters and numbers with the corresponding levels of conditions for each factor described should be established. For example, a series of letters and numbers could be used to represent not only where on the "spectrum" the patient is in terms of related criteria such as verbal and cognitive abilities and areas of strengths, but also other factors such as mental health issues, other learning differences, medical conditions, and physical disabilities. A patient's combined condition could be summarized with a series of letters and numbers such as "KEDSTVA235Z…" In addition, each combination of long codes ("KEDSTVA235Z…") should be identified as a category; for example, "KEDSTVA235Z…" could be classified as the Type 17 category. Although this approach does not intend to replace more personalized evaluations of the patient's overall conditions, a quick categorization not only makes the abstract term "on the spectrum" more concrete and gives the patient a unique profile, but it also in a way highlights all the important factors for the physician, parents, and patients to consider during the process of identifying such a code and category, ensuring that none of those important factors will be overlooked in the treatment decision process. Furthermore, once such a category code system is established, one potential benefit is that it may help parents find specific resources or communities for support. More importantly, it will make comparing cases easier in research, and more focused research could be done for different categories of patients, which will potentially help develop useful patterns in future treatment priorities.

Please note that this paper does not propose that an objective assessment should replace parents' opinions completely, but simply strongly suggests that parents should make many better-informed decisions together with the patients and the clinicians, rather than basing their decisions on what seems to bother them the most. After such an educational program and assessment tool are developed and used, clinicians should continue to take parents' support as a priority, especially given the success of interventions still often depends largely on their involvement.

Conclusion

Families and clinicians often have the best intentions in mind when dealing with ASD treatment planning. However, due to the complicated nature of neurodiversity, especially issues related to autism, the best intentions do not always translate into treatment priorities that are the best for the patient when decision makers are not aware of all the relevant factors. Therefore, it is critical for an education program to be developed to make sure that such messages are passed to them before they make the treatment choices with their clinicians. However, since parents are unlikely to be professionals in the field, an education program with an easy-to-understand language and format, as well as an appropriate length, should be the key to the success of such a program. In addition, this paper also suggests that some type of systematic evaluation of a patient's unique situation should be developed, such as one with codes and categories, so that the seemingly complicated factors can be better summarized into a simpler format, not only for the families to better understand their overall situations and get the appropriate support but also for developing future research in treatment planning for different types of patients with similar needs.

References

1. Kapp, S. K. *Autistic Community, and the Neurodiversity Movement: Stories from the Frontline*; 2020.
2. Jaarsma, P.; Welin, S. Autism as a Natural Human Variation: Reflections on the Claims of the Neurodiversity Movement. *Heal. Care Anal.* **2012**, 20 (1), 20–30. https://doi.org/10.1007/s10728-011-0169-9.
3. Bennett, E.; Heaton, P. Is Talent in Autism Spectrum Disorders Associated with a Specific Cognitive and Behavioural Phenotype? *J. Autism Dev. Disord.* **2012**, 42 (12), 2739–2753. https://doi.org/10.1007/s10803-012-1533-9.
4. Cosden, M.; Koegel, L. K.; Koegel, R. L.; Greenwell, A.; Klein, E. *Strength-Based Assessment for Children With Autism Spectrum Disorders*; **2006**.
5. Woods, R. Exploring How the Social Model of Disability Can Be Re-Invigorated for Autism: In Response to Jonathan Levitt. *Disabil. Soc.* **2017**, 32 (7), 1090–1095. https://doi.org/10.1080/09687599.2017.1328157.
6. Kimberly E. Bodnera, Diane L. Williams, C. R. E.; J., and N.; Minshew. A Comparison of Measures for Assessing the Level and Nature of Intelligence in Verbal Children and Adults with Autism Spectrum Disorder. *Res Autism Spectr Disord.* **2014**, 8 (11), 1434–1442.
7. E., F. Epidemiology of Autism and Related Conditions. Autism Pervasive Dev. *Disord.* **1998**, 32–63.
8. Aman, M.G. Treatment planning for patients with autism spectrum disorders. *J. of Clin. Psychiat.* **2005**, 66, 38.
9. Jon Baio, N. C. on B. D. and D. D. Prevalence of Autism Spectrum Disorder Among Children Aged 8 Years https://www.cdc.gov/mmwr/volumes/67/ss/ss6706a1.htm.
10. Kapp, S. K.; Gillespie-Lynch, K.; Sherman, L. E.; Hutman, T. Deficit, Difference, or Both? Autism and Neurodiversity. *Dev. Psychol.* **2013**, 49 (1), 59–71. https://doi.org/10.1037/a0028353.

DOI: 10.36838/v4i6.8

11.Rodger, S.; Braithwaite, M.; Keen, D.; Research Online, G.; Rod ger, S.; Braithwaite, M. Early Intervention for Children with Au tism: Parental Priorities Author Copyright Statement Link to Published Version Early Intervention for Children with Autism: Parental Priorities; 2004.
12.Leadbitter, K.; Buckle, K. L.; Ellis, C.; Dekker, M. Autistic Self-Advocacy and the Neurodiversity Movement: Implica tions for Autism Early Intervention Research and Practice. Front. *Psychol.* **2021**, 12 (April), 1–7. https://doi.org/10.3389/ fpsyg.2021.635690.
13.Kapp, S. K.; Steward, R.; Crane, L.; Elliott, D.; Elphick, C.; Pellicano, E.; Russell, G. 'People Should Be Allowed to Do What They like': Autistic Adults' Views and Experiences of Stimming. *Autism* **2019**, 23 (7), 1782–1792. https://doi. org/10.1177/1362361319829628.
14.Volkmar, F.; Siegel, M.; Woodbury-Smith, M.; King, B.; Mc Cracken, J.; State, M. Practice Parameter for the Assessment and Treatment of Children and Adolescents with Autism Spectrum Disorder. J. Am. Acad. Child Adolesc. *Psychiatry* **2014**, 53 (2), 237–257. https://doi.org/10.1016/j.jaac.2013.10.013.
15.Volkmar, F.; Cook, E. H.; Pomeroy, J.; Realmuto, G.; Tanguay, P.; Bernet, W.; Dunne, J. E.; Adair, M.; Arnold, V.; Beitchman, J.; Scott Benson, R.; Bukstein, O.; Kinlan, J.; McClellan, J.; Rue, D.; Shaw, J. A.; Sloan, E.; Kroeger, K. D.; Bristol-Power, M.; Fed er, J.; Filipek, P.; Harris, S.; Lord, C.; Mesibov, G.; Paul, R.; Rogers, S.; Schopler, E.; Schreibman, L.; Stone, W.; Wing, L. Practice Parameters for the Assessment and Treatment of Children, Adolescents, and Adults with Autism and Other Pervasive De velopmental Disorders. *J. Am. Acad. Child Adolesc. Psychiatry* **1999,** 38 (12 SUPPL. 1), 32S-54S. https://doi.org/10.1016/S0890-8567(99)80003-3.
16.O'Reilly, M.; Rispoli, M.; Davis, T.; Machalicek, W.; Lang, R.; Sigafoos, J.; Kang, S.; Lancioni, G.; Green, V.; Didden, R. Func tional Analysis of Challenging Behavior in Children with Autism Spectrum Disorders: A Summary of 10 Cases. *Res. Autism Spectr. Disord*. **2010**, 4 (1), 1–10. https://doi.org/10.1016/j. rasd.2009.07.001.
17.Bradley, L.; Shaw, R.; Baron-Cohen, S.; Cassidy, S. Autistic Adults' Experiences of Camouflaging and Its Perceived Impact on Mental Health. *Autism in Adulthood.* **2021**, 3 (4), 320–329. https://doi. org/10.1089/aut.2020.0071.
18.Assouline, S. G.; Foley, M.; Colangelo, N. N.; O'Brien, M. The Par adox of Giftedness and Autism The Paradox of Giftedness and Autism Packet of Information for Professionals (PIP)-Revised (2008) *The Connie Belin & Jacqueline N. Blank International Center for Gifted Education and Talent Development*; **2008**.
19.McFadden, J. Identifying and Supporting Twice-Exceptional Stu dents with Autism Spectrum Disorder (ASD). *J. Educ. Leadersh. Action* **2017**, 4 (2).
20.Cassidy, S.; Bradley, L.; Shaw, R.; Baron-Cohen, S. Risk Mark ers for Suicidality in Autistic Adults. *Mol. Autism* **2018**, 9 (1). https://doi.org/10.1186/s13229-018-0226-4.
21.Spain, D.; Happé, F. How to Optimise Cognitive Behaviour Ther apy (CBT) for People with Autism Spectrum Disorders (ASD): A Delphi Study. *J. Ration. - Emotive Cogn. - Behav. Ther.* **2020**, 38 (2), 184–208. https://doi.org/10.1007/s10942-019-00335-1.
22.Christon, L., and Myers, B. Family-Centered Care Practices in a Multidisciplinary Sample of Pediatric Professionals Providing Autism Spectrum Disorder Services in the United States. Res. *Au tism Spectr. Disord*. **2015**, 20, 47–57.
23.Taylor, B. A.; Fisher, J. Three important things to consider when starting intervention for a child diagnosed with autism. *Behavior Analysis in Practice*. **2010**, 3(2), 52.
24.Lindgren, S.; Wacker, D.; Schieltz, K.; Suess, A.; Pelzel, K.; Ko pelman, T.; Lee, J.; Romani, P.; O'Brien, M. A Randomized Controlled Trial of Functional Communication Training via Tele health for Young Children with Autism Spectrum Disorder. *J. Autism Dev. Disord*. **2020**, 50 (12), 4449–4462. https://doi. org/10.1007/s10803-020-04451-1.
25.Razjouyan, K.; Miri, M.; Zarei, M.; Farhangdoost, H.; Tavakoli, S. Study of the Effectiveness of Comprehensive, Timely, and Fami ly-Oriented Interventions in Reducing the Symptoms of Autism in Children. *J. Fam. Med. Prim. Car*e **2020**, 9 (6), 2945. https://doi. org/10.4103/jfmpc.jfmpc_947_19.
26.Kirkham, P. 'The Line between Intervention and Abuse' – Autism and Applied Behaviour Analysis. Hist. *Human Sci*. **2017**, 30 (2), 107–126. https://doi.org/10.1177/0952695117702571.
27.Kupferstein, H. Evidence of Increased PTSD Symptoms in Autis tics Exposed to Applied Behavior Analysis. *Adv. Autism* **2018**, 4 (1), 19–29. https://doi.org/10.1108/AIA-08-2017-0016.
28.Chen, J. L.; Leader, G.; Sung, C.; Leahy, M. Trends in Employ ment for Individuals with Autism Spectrum Disorder: A Review of the Research Literature. *Rev. J. Autism Dev. Disord*. **2015**, 2 (2), 115–127. https://doi.org/10.1007/s40489-014-0041-6.
29.Hadjikhani, N.; Åsberg Johnels, J.; Zürcher, N. R.; Lassalle, A.; Guillon, Q.; Hippolyte, L.; Billstedt, E.; Ward, N.; Lemonnier, E.; Gillberg, C. Look Me in the Eyes: Constraining Gaze in the Eye-Region Provokes Abnormally High Subcortical Activation in Autism. Sci. Rep. **2017**, 7 (1). https://doi.org/10.1038/s41598-017-03378-5.
30.Arwert, T. G.; Sizoo, B. B. Self-Reported Suicidality in Male and Female Adults with Autism Spectrum Disorders: Rumination and Self-Esteem. *J. Autism Dev. Disord.* **2020**, 50 (10), 3598–3605. https://doi.org/10.1007/s10803-020-04372-z.
31.Feldman, D.H. Child Prodigies: A Distinctive Form of Giftedness. *Gifted Child Quart*. **1993**, 37(4), 188-193.
32.Lai, M.-C.; Kassee, C.; Besney, R.; Bonato, S.; Hull, L.; Man dy, W.; Szatmari, P.; Ameis, S. H. Prevalence of Co-Occurring Mental Health Diagnoses in the Autism Population: A Systematic Review and Meta-Analysis.
33.Marcus LM, K. L. S. E. Working with Families. Handb. *Autism Parasite Developmental Disorder.* **1997**, 631–649.

Author

Isabelle is a high school senior and the Co-Chair of Stanford Neurodiversity Project's Network for K-12 Neurodiversity Education and Advocacy. Being neurodiverse, she has experienced the challenges directly, which has led her to co-found her high school's neurodiversity club and establish her own nonprofit to help disadvantaged children with various conditions.

IJHSR
ijhighschoolresearch.org
■ RESEARCH ARTICLE

Silver Nanowire Flexible On-skin Electrodes for Electrophysiological Monitoring

Jonathan Lu, Aidan Hong
Concordia International School Shanghai, 999 Mingyue Rd, Pudong, Shanghai, China, 201203; jonathan2023146@concordiashanghai.org

ABSTRACT: On-skin flexible electrodes are critical components of future wearables for electrophysiological (EP) signal monitoring. High electrical conductivity, flexibility, and adhesion to the skin are some of the key requirements for such electrodes. In this research project, we successfully developed a novel on-skin flexible electrode prototype by combining highly conductive silver (Ag) nanowire (NW) with highly flexible and tacky silicone adhesive film through a novel and simple transfer process. Due to the superior electrical conductivity of silver, AgNWs provide the electrode with high conductivity. The silicone adhesive film offers outstanding adhesion to skin and stretchability to ensure low skin contact impedance and comfort for users. We developed one fast and potentially scalable process to fabricate an AgNW-based flexible electrode prototype by effectively transferring a thin layer of AgNWs onto the surface of silicone adhesive film. The fabricated electrode prototype met all the electrical and mechanical requirements and demonstrated feasibility as a potential electrode for future wearable EP monitoring devices.
KEYWORDS: Materials; Nanomaterials; Silver Nanowire; Flexible On-skin Electrodes; Electrophysiological Monitoring.

■ Introduction

Most human body movements are driven by low-level electrical potentials (electrophysiological signals). Electrodes must be used to capture and monitor these electrophysiological signals. Conventional Ag/AgCl (silver/silver chloride) electrodes have many limitations including potential skin irritation caused by the gel used and poor signal quality collection during motion due to the rigid nature of the electrode.[1]

Flexible on-skin electrodes have attracted great attention in the past several years.[2-5] The electrodes can be used to monitor how well the heart's electrical signals are doing, track down abnormal heartbeats, and detect any possible problems. Much research has been done in the past years to develop various types of on-skin electrodes such as thin metal film electrodes, nano carbon (such as carbon nanotube) based electrodes, and metallic nanomaterial-based electrodes.[6] But all these electrodes require complicated fabrication processes such as metal sputtering, use of organic solvents, etc.

This research aimed to create an on-skin electrode that will be able to replace conventional Ag/AgCl electrodes using silver nanowire through a simple and effective transfer process. The electrode must have high conductivity to ensure that the data and accuracy of the testing are maximized. The on-skin electrodes also must be comfortable and have a stretchability of more than 15% which is the stretchability of human skin.[6]

Essentially, the objective of the project was to use silver nanowire to develop a soft, self-adhering, stretchable, and conductive on-skin electrode for electrophysiological monitoring and human-machine interfaces.

■ Methods

Materials:

Silver nanowire (45nm in diameter and 20μm in length) isopropanol (AgNW-IPA) solution (0.67 weight %) was from the Shenzhen Institute of Electronic Materials (SIEM). The filtration paper (125 mm diameter) was from Whatman. The silicone adhesive is a standard 2-component heat-curable silicone adhesive from Henkel. The water-soluble polyvinyl alcohol (PVA) film (35U) is from Suzhou Jiang Sheng Garment Accessories. The glass slides (76x22x2 mm) were from Sangon Biotech. The copper foil tape was from 3M China.

Apparatus:

The spray bottle (SLF-22) was from Cangzhou Victory Medical Packaging. LCR meter (VC4091A) was from Shenzhen Victor instrument Co. Ltd. ECG (Electrocardiogram) lab tester was from Fudan University. The tensile tester (68TM-5) with AVE2 and surface tackiness tester (KJ-6032) were from Instron and Kejian Instruments, respectively.

AgNW Electrode Fabrication Process:

Three different processes were explored to fabricate the AgNW on-skin electrodes. The first method was to directly spray the AgNW-IPA from a spray bottle onto a flat piece of polyvinyl alcohol film (PVA) which is water-soluble. The spraying was repeated 5 times to ensure there were enough nanowires to form a continuous layer. A sufficient time (roughly one minute) was given to allow the IPA to dry out before the next spraying. A liquid silicone adhesive was then dispensed onto the surface of the PVA with the AgNW layer facing up. After the liquid silicone adhesive was solidified in a heating oven for 10 minutes at 100˚C, a silicone-AgNW-PVA stack was then put into hot water (90˚C) for 6 hours to dissolve the PVA layer and expose the silver nanowires on top of the silicon film.

The second method was to spray the AgNW-IPA directly onto the glass slides which were pre-cleaned by IPA. The spraying was repeated 5 times to ensure there were enough nanowires to form a continuous layer. The liquid silicone was then dispensed to the glass slide with the AgNW facing up and cured the same way as in the first process. Then the silicone film was detached from the glass slide very gently.

DOI: 10.36838/v4i6.9

The third method (as shown in Figure 1) was to spray the AgNW-IPA directly onto a filter paper with a mask which was used to define the final AgNW layer pattern. The spraying was repeated 5 times. After the masks were removed, a piece of pre-solidified silicone film (roughly 30mm *30mm*2mm) was placed firmly onto the AgNW pattern on the filter paper. The silicone film was gently hand-pressed for a few seconds to ensure uniform contact between the AgNW layer and the silicon film surface. Then the silicone film was slowly and carefully peeled off to transfer the AgNW pattern off the filter paper.

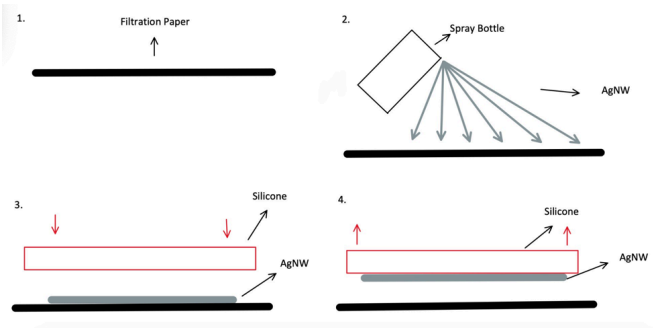


Figure 1: Schematic process flow of AgNW transfer.

Results and Discussion

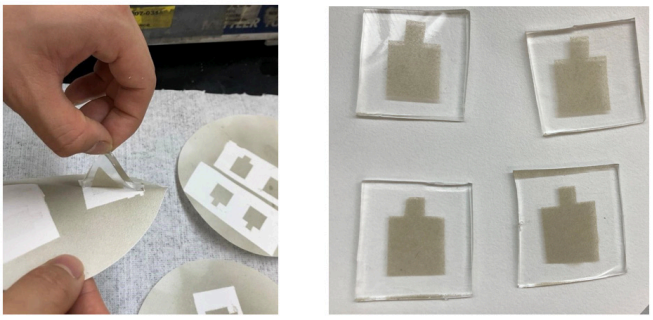
AgNW Electrode Prototype Fabrication:

The first method was tested in which a water-soluble PVA film (~40μm thick) was used to transfer the nanowire. When PVA film is dissolved in hot water, the transfer of the AgNW layer on the silicon film surface should be exposed. Surprisingly, even though the PVA film was dissolved by the hot water, the formed electrode (silicone film/AgNW layer) did not show high surface electrical conductivity (the measured surface resistance was on the level of mega ohms).

The nanowire was sprayed directly onto the glass slides in the second testing method. Similar to the first method, the silicone film/AgNW layer after being detached from glass slides did not show high surface electrical conductivity (the measured surface resistance was on the mega ohms level too). In addition, detaching silicone film from the glass turned out to be quite challenging due to the high tackiness of the silicone adhesive film.

Due to the extremely high surface reactivity of nanowires, the nanowires are preferentially covered by polymeric resins. We believed that the liquid silicone probably fully covered and embedded the nanowires and left minimum nanowires exposed on the silicone surface, causing low surface electrical conductivity.

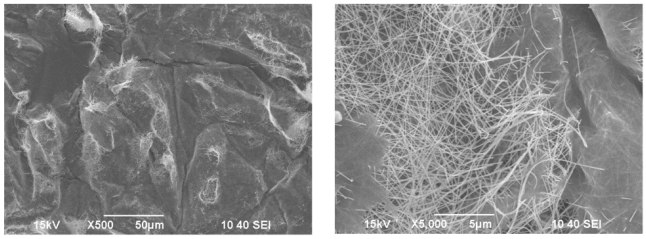
In the third process, the nanowire solution was sprayed directly onto filtration papers with masks which were used to define the patterns of the final electrodes. Unlike the first two methods, this method turned out to be very effective. Due to the porous nature of the filter paper, detaching or peeling off the silicone film from the filtration paper was very easy, and the majority of the nanowire stayed with the silicone film (Figure 2a). The fabricated AgNW electrode prototypes are about 30mm x 30mm x 2mm in dimensions (Figure 2b).



(a) (b)

Figure 2: AgNW transfer process (a) and fabricated AgNW electrode prototypes (b).

The SEM (scanning electron microscope), SU3800 from Hitachi, was employed to observe the silver nanowires on the electrodes. From SEM photos in Figure 3(a), the nanowires formed a continuous layer on the surface of the silicone film. From Figure 3(b), it can be seen that the nanowires partially penetrate the silicon film surface, which will ensure the nanowires will not fall off during future use. More future work should be conducted to study the correlation between surface tackiness of silicone and AgNW penetration into the silicone surface so that the AgNW penetration could be better controlled.



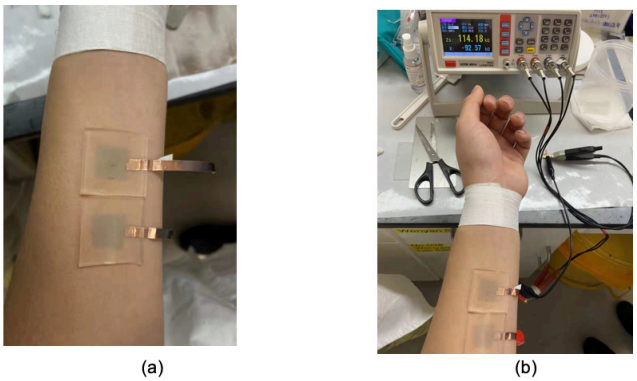
(a) (b)

Figure 3: SEM photos of the transferred AgNWs on the surface of the electrodes at 500 magnification (a) and 5000 magnification (b).

Electrical and Mechanical Performance of AgNW Electrode Prototype:

The electrical impedance of the AgNW electrode prototype was measured by placing two pieces of copper foil tape (30mm apart), which act as the connecting leads for the LCR meter, on the forearm skin, which was cleaned using alcohol wipes, and then aligning and attaching two AgNW electrodes onto the forearm (Figure 4).

The electrical impedance under frequencies ranging from 100Hz to 10,000Hz was collected for the AgNW electrode prototype and the standard Ag/AgCl electrode (as shown in Figure 5). As seen in the figure, the electrical impedance of the developed AgNW electrode is similar to or slightly lower than that of the standard electrode. This shows that the electrical performance of the developed AgNW electrode prototype is acceptable.



(a) (b)

Figure 4: Electrical Impedance Testing of AgNW Electrodes.

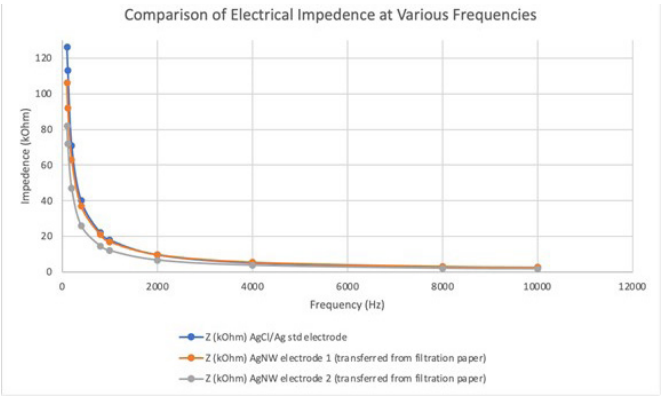
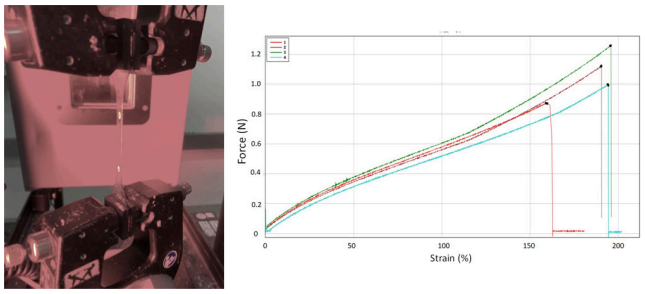


Figure 5: Electrical impedance (Z) with frequency for conventional Ag/AgCl electrode and AgNW electrodes.

Mechanical Performance of the AgNW Electrode Prototypes:

The elongation of the silicone material was tested by following ASTM D638.[7] The silicone film was cut into dog bone shape with specific dimensions, and it was tested using an Instron tester (Figure 6a). Figure 6b is the force-strain data collected from four specimens. As seen in Table 1, the average elongation (stretchability) of the silicone material is around 184.95%, which well exceeds the 15% elongation requirement (Figure 6; Table 1).



(a) (b)

Figure 6: Dog-done shaped silicone film on an Instron tensile tester (a) and force and strain data of the silicone material (b).

Table 1. Tensile strength and elongation of the silicone material.

Length (mm)	Thickness (mm)	Tensile Strength at Maximum Force (MPa)	Elongation at Break (%)
6.3	2	0.07	159.82
6.3	2	0.09	190.29
6.3	2	0.1	195.42
6.3	2	0.08	194.28
Average			184.95

The surface tackiness of the silicone film was studied using rolling ball surface tack testing (comply with GB/T 4852-2002)[8] where a small stainless-steel ball was placed on the top of the piece of silicone film which rests on a platform with a 30-degree slope. It was found that the silicon film has a surface tackiness of ball 12 at a 30-degree slope which is very high compared to those of other regular adhesive materials (Figure 7).

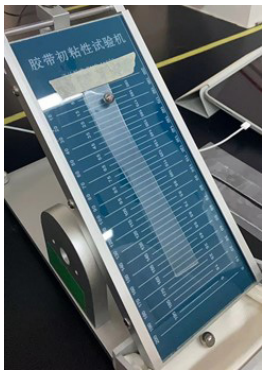


Figure 7: Surface tackiness testing of silicone film.

Feasibility Demonstration of the AgNW Electrode Prototypes:

A lab electrocardiogram (ECG) setup was used to demonstrate the feasibility of the AgNW electrode prototype we developed. Two AgNW electrodes were attached to two forearms, and one was placed on the skin of one side of the waist as the reference electrode.

The collected ECG graph below (Figure 8) shows a clear rhythm and a relatively high signal-to-noise ratio (SNR). This study indicates that the developed AgNW electrode prototypes are acceptable for collecting ECG data and could be a potential electrode for future wearable electrophysiological monitoring.

(a) (b)

Figure 8: ECG testing set up (a) and graph data collected using the AgNW electrode prototype (b).

Conclusion

We developed a simple and novel process to effectively transfer AgNW onto a highly tacky silicone surface to form a self-adhering flexible on-skin electrode. The electrical and mechanical performance all meet or exceed the requirements in comparison to conventional Ag/AgCl electrodes. The proof of the concept of using fabricated AgNW electrode

DOI: 10.36838/v4i6.9

prototypes for electrophysiological monitoring was confirmed by the lab-scale ECG testing.

Acknowledgements

The authors would like to thank Mrs. Anne Love to be a sponsor for this project, Dr. Daniel Lu from Henkel being the adult supervisor, and Prof Zhuo Li from Fudan University for her help with the ECG testing.

References

1. Y. Ma, Y. Zhang, S. Cai, Z. Han, X. Liu, F. Wang, Y. Cao, Z. Wang, H. Li, Y. Chen, X. Feng, Adv. Mater. 2020, 32, 1902062.
2. C. Wang, C. Wang, Z. Huang, S. Xu, Adv. Mater. 2018, 30, 1801368.
3. Y. Liu, K. He, G. Chen, W. R. Leow, X. Chen, Chem. Rev. 2017, 117, 12893.
4. S. P. Lacour, G. Courtine, J. Guck, Nat. Rev. Mater. 2016, 1, 16063.
5. J. C. Yang, J. Mun, S. Y. Kwon, S. Park, Z. Bao, S. Park, Adv. Mater. 2019, 31, 1904765.
6. H. Wu, G.G. Yang, K.H. Zhu, S.Y. Liu, W. Guo, Z. Jiang, and Z. Li, Adv. Sci. 2020, 2001938
7. https://www.instron.cn/zh-cn/testing-solutions/by-standard/astm/astm-d638
8. https://www.chinesestandard.net/PDF.aspx/GBT4852-2002

Author

Jonathan Lu and Aidan Hong both are 11th-grade students at Concordia International School Shanghai in Shanghai China.

IJHSR
ijhighschoolresearch.org
REVIEW ARTICLE

A Systematic Review of Amyloid-β and Tau's contribution to Neuroinflammation in Alzheimer's Disease Progression

Avni Kabra
East Lyme High School, 30 Chesterfield Rd, East Lyme, CT, 06333, USA; avnik024@gmail.com

ABSTRACT: Alzheimer's Disease (AD) is classified as the most common neurodegenerative disease that triggers and furthers the onset of neuronal cell death as a consequence of hallmark pathological changes. Since there is yet no cure for dementias such as Alzheimer's Disease, there is perpetual cognitive function and developmental decline in diseased patients. The hallmarks of AD include abnormal amyloid-β plaques, neurofibrillary tangles, and neuroinflammation. While these neuronal processes are interlinked, it is not entirely understood how they are related or regulate each other. This review will focus on the association of neuronal processes and Alzheimer's Disease pathology - more specifically the influence of amyloid-β peptide aggregation on tau tangles, and microglial inflammation. In this paper, the interaction between amyloid-β and tau tangles including how this interaction leads to proinflammatory patterns, which results in neuroinflammation observed in Alzheimer's Disease, was analyzed. In context, evidence for induction and progression of tau hyperphosphorylation and neuroinflammation were observed through analysis of iPSC neurons. In examining the pathological hallmarks for Alzheimer's Disease, neuronal abnormalities and dysfunctions have been analyzed with the intent of uncovering new potential therapeutic approaches for rescuing degenerative cell function and structure.

KEYWORDS: Cellular and Molecular Biology; Neurobiology; Alzheimer's Disease; Hallmarks.

Introduction

Alzheimer's Disease (AD) is a specific dementia that accounts for 60–80% of all dementia cases and affects 33.9 million people worldwide - a number that is expected to triple in the next 40 years.[1] AD is characterized as loss of cognitive function, such as thinking, remembering, reasoning, and behavioral abilities that affects patients' daily lives. Late-onset Alzheimer's Disease (LOAD), the most common form, is present with an onset of ~65 years of age, while early-onset Alzheimer's Disease (EOAD) is present in the younger population, with the age of onset between 35 through 60 years of age.[2] Although the age of onset can vary, according to the current criteria, the process to diagnose a patient is common and widely accepted. Diagnosis is only determined with varying degrees of certainty as possible or probable Alzheimer's Disease when all other causes have been excluded. However, currently, an unambiguous detection of AD depends on verification of typical Alzheimer's Disease pathological changes in brain tissue by autopsy reports and post-mortem analysis. As the disease progresses, patients develop increasingly severe disabilities and become completely dependent on patient care at later stages of disease.[3] The accumulation of hallmark AD pathology, first identified in the medial temporal lobe and hippocampus regions, ultimately triggers the clinical symptoms of dementia and short-term memory loss. Early stages of dementia also include subtle cognitive impairment and health issues, such as anxiety and delusions, which then develop in the lateral and parietal lobe, furthering cognitive impairment. These include failing to recall personal history, poor recognition and direction sense, and withdrawing socially. As it proliferates into the frontal lobe, occipital lobe, and cerebellum, patients fail to perform rudimentary skills, lose vision and sense of direction, and become unable to communicate.[3,4] This progression is due to cell death, the loss of neurons and neuronal connections, which is influenced by the pathological hallmarks of Alzheimer's Disease. These include amyloid-β peptide aggregation, neurofibrillary tangles, and neuroinflammation. Investigating and targeting the primary cause of Alzheimer's Disease may provide a foundation to tackle a defective component before it damages the nervous system.

Amyloid-β and Alzheimer's Disease Pathology:

A prominent hallmark pathology and primary indicator of Alzheimer's disease is the extracellular plaque deposits of the β-amyloid peptide. As suggested by the Amyloid-Cascade hypothesis, β-amyloid is the main cause of Alzheimer's and the misfolding of the extracellular Aβ protein accumulated in senile plaques causes memory loss and confusion, resulting in personality changes and cognitive decline over time.[5] Dysfunction in the mechanisms of Aβ production, specifically APP (β-amyloid precursor protein) proteolysis, has been linked to plaque formation.[6] APP is a type-I oriented membrane protein that plays a vital role in biological transport, such as neuronal development, signaling, and intracellular transport.[7] APP is processed through two principal pathways: the amyloidogenic pathway and the non-amyloidogenic pathway.[8] In the non-amyloidogenic pathway, APP is cleaved by the protease enzyme α-secretase. This processing produces a partial APP CTF, which lacks the amino terminal of the Aβ domain, thus creating a non-toxic byproduct. On the other hand, the amyloidogenic pathway produces the two forms of the amyloid-β peptide, Aβ40 and Aβ42, through cleavage of APP using β-secretase and γ-secretase.[9] A defect in the APP processing system or in the APP itself will yield accumulation

DOI: 10.36838/v4i6.10

of amyloid-β. A mutation in APP, such as the duplication of chromosome 21 or duplication of the APP gene, accelerates generation of amyloid-β after being cleaved by β-secretase 1 (BACE1) and γ-secretase, therefore, furthering the progression of Alzheimer's Disease at a younger age. Similarly, a mutation in the genes that constitute the γ-secretase catalytic subunit, presenilin 1 (PSEN1) and presenilin 2 (PSEN2), alters gamma secretase function which leads to dysfunction in the proteolysis of APP and production of amyloid-β.[7] Whereas Aβ40 accounts for approximately 90% of amyloid-β production, the Aβ42 product is more hydrophobic and is suggested to be involved in amyloid-β aggregation when mutated, leading to amyloid plaque deposition in the AD brain.[9,10] As accepted and useful as the Amyloid Cascade Hypothesis may seem, there are growing amounts of research findings that contradict this theory. Current research suggests that individuals with substantial plaque aggregation can have normal cognition.[11] Research has also shown after Alzheimer's Disease onset, immunoclearing of senile plaques may not always improve cognition, although Alzheimer's Disease itself is based on plaques and plaque deposits.[5] This seems to indicate that amyloid-β deposits alone are not always sufficient to cause dementia.[11-13]

Neurofibrillary Tangles and Alzheimer's Disease Pathology:

Neuropathological studies of Alzheimer's Disease link clinical symptoms to accumulation of neurofibrillary tangles (NFTs), which are constituted by aggregates of hyperphosphorylated tau protein. Abnormal NFTs accumulation occurs within the neuronal cell bodies, including limbic and association cortices, which undergo degeneration during the course of AD.[14] These are present in the regions responsible for various cognitive domains that are impaired in AD, with the density of tau correlating with the region experiencing cognitive decline.[15] Each of these fibrillary lesions contains microtubule associated protein tau (MAPT), which is characterized as a paired helical filament structure (PHFs), and in a decreased amount, straight filaments.[16] Neuronal tau assists tubule networking by assembling and stabilizing microtubules through binding to microtubule filaments, in turn improving cytoskeleton structure and allowing neurite extension.[17] Tau is modified post-translationally through phosphorylation, which inhibits its binding to microtubules and could ultimately lead to tau hyperphosphorylation when all potential sites are fully phosphorylated. Upregulated phosphorylation at specific microtubule associated proteins (MAP), however, promotes tau self-aggregation into filaments.[18] Thus, tau hyperphosphorylation impairs microtubular structure and tau itself becomes destabilized. This destabilization generates dystrophic neurites, triggers formation of neurofibrillary tangles and neuropil threads, which are enlarged filament-containing dendrites, axons, and terminals.[18,19] Abnormally hyperphosphorylated tau, a key component of PHFs in Alzheimer's Disease, gains a toxic ability to sequester regular tau and other MAPs, causing microtubule disassembly.[18] NFTs, as the most mature form of hyperphosphorylated tau, leads to neurofibrillary degeneration, which degrades neurons and dendrites to influence the onset of AD.[19]

Neuroinflammation and Alzheimer's Disease Pathology:

Alongside amyloid-β plaques and neurofibrillary tangles, neuroinflammation plays a significant role in AD pathology. The term "neuroinflammation" suggests an inflammatory response in the central nervous system (CNS) when there is an insufficiency in eliminating the foreign intruder. Inflammation is, therefore, intended for neuroprotection, but overexpression promotes tissue damage and neurodegeneration pathology.[20] Neuroinflammation was originally observed near aggregated amyloid-β peptides in autopsy reports of neuropathological disorders, including AD.[21] Glial cells in the CNS, specifically astrocytes and microglia, are involved in the inflammatory response. Astrocytes are a type of macroglia that are vital for mediating glutamate levels, providing structural support for the blood brain barrier (BBB), and releasing pro-inflammatory cytokines, such as TNFα, to moderate extravasation from blood to the CNS parenchyma. These reactive astrocytes lead to degradation of synapses.[22] Microglia are macrophages that are vital for neuronal plasticity, neurogenesis, regeneration, and defense. They have the potential to phagocytose neurotoxic components, secrete cytotoxic entities, and produce and provide antigens.[23] In the absence of invasive pathogens, microglia remain in an inactive state during which they scan and monitor the parenchyma of the brain sporadically, without disruption of neuronal activity and other glial cells. When any potential damage is detected, microglia activate and transform into mobile amoeboids in order to reach the defective module.[24] These glial cells release cytokines and neurotoxins for a sustained period of time, sometimes up to hours, further worsening pathology in the CNS.[25] Neurotrophic factors secreted by activated neurons, cytokines, are linked to inhibition of the expression of Class II Major Histocompatibility Complex (MHC II) in regards to microglial cells.[26] Thus, the anti-inflammatory phenotype transforms into a pro-inflamsmatory combative phenotype. This stimulation begins a chronic inflammatory response, resulting in behavioral changes, neuronal death, and rapid progression of Alzheimer's Disease.[27]

■ Discussion

Amyloid-β and Tau During Alzheimer's Disease Progression:

To investigate the interaction between tau and amyloid-β, it is relevant to understand the effects of hyperphosphorylated tau and amyloid-β during Alzheimer's Disease progression. Chen and colleagues performed an analysis on 82 patients in the following treatment groups: mild cognitive impairment (MCI), Alzheimer's Disease, and healthy controls.[27] This model displayed significant changes in amyloid-β biomarkers, Aβ42 and Aβ40, and tau biomarkers t-tau (total tau) and p-tau (phosphorylated tau) over the course of a 3-year study period. T-tau and p-tau are suggested to be released during neuronal degradation, with p-tau correlating with NFTs and t-tau correlating with impaired axonal transport. While the Aβ42 levels represented through Aβ42-related biomarkers changed as a function of time in patients with stable AD, tau-related biomarkers changed linearly as a function of clinical diagnostic classifications over the AD cognitive spectrum. These findings associate amyloid-β and tau with ongoing indepen-

dent progression in etiological processes.[28] It is important to recognize that the trials that were done cannot quantify levels of amyloid-β and tau, and thus, only correlations and associations can be observed. Growing research suggests that Aβ and tau change in a sequential, dynamic and temporarily overlapping manner.[29,30] Furthermore, a neuropsychological cohort study with a MCI group showed, at baseline, a negative correlation of Aβ42 with global cognition. In the second year, Aβ42 and phosphorylated tau correlated negatively with memory, function, and cognition, while in the third year the amyloid-β biomarkers displayed positive correlation with episodic memory, semantic memory, and attention, and Aβ42 and hyperphosphorylated tau represented a positive correlation with all cognitive functions. Therefore, there is negative association of Aβ42 with cognitive function during the first- and second-year trials, and a positive association in the third trial. This displays that cognitive impairment in the early stages better fits with hyperphosphorylated tau increase than with amyloid-β. Yet, amyloid-β is more expressive in the Alzheimer's Disease group (including stable AD subjects and MCI that progressed to AD subjects).[28] Given the strong correlation between hyperphosphorylated tau accumulation, rather than amyloid-β, and early stages of AD, it is unclear whether the overlapping processes and hyperphosphorylated tau interactions within the brain parenchyma and destabilization of the cytoskeleton caused further expression in amyloid-β seen in Stage 3 of the study. Altogether, this data suggests that hyperphosphorylated tau is a better indicator of Alzheimer's Disease, and there are possible implications on amyloid-β production. It is important to note that Chen and collaborators were measuring biomarkers from blood samples, and researchers are currently uncertain of how hallmarks, such as amyloid-β aggregation and tau, correlate with AD in plasma.

To confirm and further explore the idea of tau and amyloid-β interactions, Barthélemy and colleagues analyzed tau and amyloid-β pathophysiology throughout AD development, which was measured using clinical assessments, cerebral spinal fluid (CSF), and brain imaging techniques. Through observations of a study cohort consisting of non-familial at-risk and symptomatic participants, tau isoforms were compared to the stages and proliferation of amyloid plaques using mass spectrometry methods. When amyloid plaques are initiated, there is an increase in levels of tau phosphorylation sites, pT217 and pT181. As neuronal degradation begins to occur, dictated upon decreased cortical metabolic activity, levels of the pT205 phosphorylation site and tau isoform, t-tau, begin to increase as well. Based on clinical and cortical atrophy decline, tau-PET tangles (NFTs) evolve, although pT217 and pT181 decrease. It is evident that tau sequentially changes with disease stage, amyloid plaques, and cortical atrophy. This confirms that the hyperphosphorylation of tau is dynamic and, in early stages, depends on the proportion of mutated amyloid-β present, and later, on the degree of cognitive decline. This also demonstrates tau isoforms are contingent upon aggregated amyloid plaques, therefore providing a possible link between amyloid-β and tau. Additionally, given the predictability of tau as stages progress, abnormal production or alterations to tau may be used to identify early stages of AD. Then, to assess whether amyloid-β aggregation in varying cortices is associated with tau phosphorylation, Barthélemy and colleagues used cross-sectional correlations, based on bivariate regressions, between phosphorylation sites for p-tau and cortical regions of amyloid plaque accumulation through PiB-PET SUVR, which is a protein that binds to amyloid-β and can be detected through PET imaging. The results demonstrate a positive correlation between pT217/T217, pT181/T181 and pT205/T205 phosphorylation and PiBPET SUVR, although pS202/S202 was negatively correlated with the amyloid plaque biomarker.[33] This data signifies that although tau is the most accurate marker for AD, amyloid-β initiation occurred first. The variable levels of correlation further the hypothesis of tau influencing aggregation of amyloid-β, with different cortical metabolic activities and presence triggering positive and negative correlations.

Effect of Tau on Amyloid-β Pathology:

To understand if there is an effect of tau expression on amyloid-β production, Pickett and collaborators performed recent research on pathology and behavior in context of AD using transgenic mice models APP/PS1+Tau mice and three control genotypes: control (MAPT -/-), APP/PS1 only (MAPTnull x APP/PS1), and human tau only (MAPTnull x rTg21221). APP/PS1+Tau mice contained APP, PSEN1, and Tau, while the MAPT -/- mice lacked the MAPT gene due to double knockout. In APP/PS1+Tau mice and APP/PS1 mice, onset of amyloid senile plaques appeared in cortex and hippocampus after six months, with plaque burden positively correlating with age. Plaque deposition was observed through the pan-Aß antibody and stained with 0.05% thioflavine S in 50% ethanol.[31] Remarkably, lower plaque deposits were identified in APP/PS1+Tau mice, rather than in APP/PS1 mice, which is unusual compared to previous studies in which larger plaques were identified.[32] Although both mice were expressing the APP/PS1 transgene, senile plaque deposits differed. In a post-mortem analysis of mice of the same cohort that were being tested, percentage accumulation of amyloid-β in pre- and post-synaptic terminals in APP/PS1 and APP/PS1+Tau mice were higher than the controls, although the amyloid-β levels in these mice showed similar percentages of amyloid-β concentration. There was also no effect on amyloid-β aggregates in synapses when lowering tau levels. Tau was also not detected in mice that didn't express tau (controls and APP/PS1). Furthermore, in contrast with amyloid-β, the percentage of tau near synapses did not vary significantly, but the number of synapses with tau notably changed.[31] Damaged synapses are a strong pathological indication and driver of cognitive decline within Alzheimer's Disease.[30,32] These findings suggest that tau and amyloid-β was observed in both pre and post synapses and supports the hypothesis of the interaction of tau and amyloid-β in dysfunction and cognitive impairment. Yet, decreasing hyperphosphorylated tau displayed no effect on amyloid-β generation, implying that hyperphosphorylated tau does not factor in on mutated amyloid-β, although lower plaque deposits were identified in mice with tau (MAPTnull x rTg

DOI: 10.36838/v4i6.10

21221). This suggests that hyperphosphorylated tau does not affect amyloid-β initiation and progression. It also suggests the possibility of either tau function in mice varying from the function of tau in the human nervous system, or the temporal and spatial expression generated through Ca2+/calmodulin-dependent protein inaccurately depicting the endogenous tau. Additionally, deregulation of tau levels in APP/PS1+Tau mice did not alter amyloid-β aggregation. Yet, APP/PS1+Tau mice contained decreased amyloid plaque depositions in comparison to APP/PS1 mice, thus suggesting that while stabilized tau does not impact amyloid plaque levels, tau after hyperphosphorylation influences amyloid-β progression.

Effect of Neuroinflammation on Amyloid-β Pathology:

The first signs of neuroinflammation stimulating Alzheimer's Disease was noticed when microglia upregulation was seen in post-mortem analysis of patients with neuropathological disorders.[34-36] These early studies observed microglial activation near amyloid-β aggregation sites, leading them to believe neuroinflammation amplified cognitive decline and onset of AD.[21] Multiple studies have already determined that anti-inflammatory microglia promote Aβ clearance to halt progression of AD, although continuous microglial accumulation will release cytotoxins, such as proinflammatory cytokines, reactive oxygen species (ROS), chemokine, and complement proteins. Microglia and astrocytes will then experience loss of function and BBB structural integrity, and the toxic molecules secreted will upregulate Aβ production, while a lowered expression of degradative enzymes decreases Aβ clearance.[37]

To further analyze the effects of neuroinflammation on amyloid-β pathology in AD, Marsh and colleagues backcrossed 5xfAD mice onto a Rag2-/-/Il2rγ-/- double-knockout to create mice that lacked immune system cells, specifically, T cells, B cells, and natural killer (NK) cells. These mice, Rag-5xfAD, and wild-type mice, Rag-WT, were corresponded with immune-competent Alzheimer's Disease transgenic and wild-type mice, WT-5xfAD and WT-WT. Then they performed a highly sensitive multiplex ELISA, a protein measurement technique, to detect amyloid-β levels. Marsh and colleagues found that all amyloid-β species in Rag-5xfAD half-brains were increased twofold in comparison to WT-5xfAD half-brains. Rag-5xfAD mice also displayed a highly significant increase of plaque volume. Marsh and collaborators then had to determine if this substantial increase of amyloid-β was prompted by increased Aβ production or decreased clearance. Thus, they analyzed a Western blot of the protein level of APP and Presenilin 1. Rag-5xfAD and WT-5XfAD mice displayed an expected increase in APP and Presenilin 1 in comparison to wild-type controls, yet there was no difference in protein expression between Rag-5xfAD and WT-5xfAD. Then, they performed a hierarchical cluster analysis of microglial- enriched genes that were recognized from a recently published online RNA-seq database. Those genes strongly implied microglial dysfunction, which brings up the question if alterations in microglial number and formation could be the reason for increased amyloid peptides. Using immunohistochemical and unbiased IMARIS 3D rendering of microglia, Marsh and colleagues were able to quantify and model the microglial process. Their findings were that WT-5xfAD mice displayed increased microglia cells and decreased microglial branching processing in comparison to WT-WT mice, suggesting an activated phenotype. Additionally, Rag-5xfAD mice displayed a substantial increase in microglial number and decreased branching and process length when compared to WT-5xfAD mice. Surprisingly, Rag-WT microglia had a longer process and a smaller increase in branching on average, when compared to WT-WT microglia.[39] All this data suggests that the loss of immune cells modulates microglial phenotypes and processes. Differentiation between Rag-5xfAD and WT-5xfAD suggests that microglia in Rag-5xfAD mice are further or alternatively activated. It also appears that the observed elevations of senile plaque load in Rag-5xfAD are not due to increased or abnormal APP production and/or processing, but rather are likely mediated because of altered Aβ clearance, which is the primary function of microglia. Thus, it can be concluded that defective microglial activation and processing leads to aggregation of amyloid plaques.

Amyloid-β and Tau with Neuroinflammation

It is crucial to investigate the correlation between tau, amyloid-β, and pro-inflammatory patterns regarding the stages of AD pathology. To do so, Dani and colleagues used a cohort study of 26 patients with varying degrees of Alzheimer's Disease and mild cognitive impairment (MCI). They examined PET imaging of tau, amyloid-β, and microglia using biomarkers, 18F-AV-1451 (tau biomarker) and 11C-PK11195 (microglia biomarker) and analyzed it through Z-score mapping. Z-score mapping is used to statistically compare data points by measuring standard deviations above or below the mean. As expected, in amyloid-positive individuals, there was a high positive correlation with microglia, and higher correlation coefficients were identified in the MCI group, rather than the AD group. When those same individuals were tested for tau correlations, there were strong positive correlations with microglia, especially in the temporal lobe of AD subjects. Then, they observed tau in amyloid-negative individuals (AD without appearance of senile plaques). Tau aggregation and microglial activation were positively correlated, yet the correlation coefficients were smaller and the Z-score was lower in comparison to amyloid-positive patients. Microglial activation, amyloid-β deposition, and tau aggregation displayed a strong positive correlation in the temporal lobe.[38] This data supports amyloid-β uptake by microglia as being a probable stimulus of pro-inflammatory patterns. The positive correlation of microglia and tau also suggests an association between the two processes, although what is specifically being influenced cannot be determined without further testing and a meta-cohort study. It is important to recognize, however, that the strongest positive correlation between the two was seen in the temporal lobe, which is where tau aggregation increases aggressively during the Braak stages. This could suggest tau's rapid increase is due to neuroinflammatory patterns, which is influenced by aggre-

gation of amyloid-β in early stages, or that neuroinflammatory patterns occur because of the sudden aggregation of tau and amyloid-β exacerbates the effect.

Given microglial activation is stimulated by amyloid-β, it is important to analyze the correlation between neuroinflammation and amyloid-β and tau pathology to understand AD. Prokop and colleagues recently established a cohort of 66 patients with varying stages of Alzheimer's Disease progression, primary age-related tauopathy (PART), and pathological aging conditions. They initially analyzed tau's effect on microglia activation with insignificant deposits of amyloid-β in the temporal lobe. Although PART patients displayed high levels of NFTs and neuropil threads with no simultaneous amyloid-β progression, there was a negligible increase of activated microglia in AD patients. On the other hand, patients with severe AD displayed high levels of plaques, tau, and activated microglia in respect to the controls. Then, Prokop and collaborators analyzed neuroinflammation in twelve pathological aging cases, or those with the presence of senile plaques and nearly no tau indications. Comparatively, microglial activation in PART cases was still significantly low, and severe AD cases again showed a major increase in neuroinflammation. Yet it is important to consider that in both groups, subjects presented overall higher averages than the controls, suggesting amyloid-β and tau pathology is not exclusively due to ageing. Since amyloid-β pathology appeared to have a significant influence on microglia activation, it was also observed that higher microglia levels related with amyloid-β plaques through PHF-1 immunohistochemical staining. Moreover, a detailed computer-assisted image analysis was performed to correlate the percentage of area covered by plaque deposits and progression of anatomical progression of AD, in which there was an increased aggregation of amyloid-β deposits. Although, the area covered by hyperphosphorylated tau increased with disease progression, there were comparatively lower amounts of hyperphosphorylated tau observed in all brain cortices.[40] This data shows that tau pathology alone cannot induce substantial pro-inflammatory patterns, and concurrent amyloid-β is necessary. Similarly, this data suggests it can be confirmed that amyloid-β pathology alone can trigger severe neuroinflammation, with tau exacerbating that effect.

Conclusion

Alzheimer's Disease is recognized by the exacerbation of amyloid-β peptides, hyperphosphorylated tau, and neuroinflammation. Over the years, studies have associated these hallmark pathologies, although the exact interactions are unclear. Analysis of AD pathology literature suggests amyloid-β has a strong correlation with tau hyperphosphorylation, although a meta-analysis must be done to confirm findings of studies in AD patients. Tau, which becomes hyperphosphorylated, is the greatest and most influential on cognitive decline. Cognitive decline translates to a rapid progression of AD pathology. Amyloid-β, additionally, has the potential to provoke pro-inflammatory patterns. Tau, however, cannot trigger neuroinflammation alone, and co-occurrent amyloid-β pathology is necessary for substantial microglia activation. Inflamed microglia can lead to cognitive decline, which will lead to a more rapid progression of AD. Additionally, pro-inflammatory patterns display causation of the spread of tau, similar to the pathology seen in Braak stages. Yet, a thorough analysis must be done to fully confirm neuroinflammation can cause hyperphosphorylation of tau, which then leads to cognitive decline that furthers the development of AD. These interactions shed light on progression of Alzheimer's Disease, and specifically targeting these hallmarks using multiple strategies to downregulate all other hallmarks will provide potential therapies. Future studies should specifically aim to target tau phosphorylation sites, to decelerate production of abnormal tau, while concurrently targeting APP and its processing to inhibit mutated amyloid-β. This literature review analyzed and suggested how specific pathological hallmarks interact within the brain to worsen Alzheimer's Disease symptoms. There will likely be more potential in therapies if they target two pathological hallmarks, instead of only focusing on one. This also narrows the possible pathways to test and expands current knowledge about AD pathology.

Acknowledgements

I would like to thank Dr. Marta Neves Madureira for her assistance in this process.

References

1. Barnes, D. E., & Yaffe, K. (2011). The projected effect of risk factor reduction on Alzheimer's disease prevalence. The Lancet Neurology, 10(9), 819–828. https://doi.org/10.1016/S1474- 4422(11)70072-2
2. Tellechea, P., Pujol, N., Esteve-Belloch, P., Echeveste, B., García-Eulate, M. R., Arbizu, J., & Riverol, M. (2018). Early- and late-on set Alzheimer disease: Are they the same entity? Neurologia (Barcelona, Spain), 33(4), 244–253. https://doi.org/10.1016/j.nrl.2015.08.002
3. López-de-Ipiña, K., Alonso, J.-B., Travieso, C. M., Solé-Casals, J., Egiraun, H., Faundez-Zanuy, M., Ezeiza, A., Barroso, N., Ecay-Torres, M., Martinez-Lage, P., & Martinez de Lizardui, U. (2013). On the selection of non-invasive methods based on speech analysis oriented to automatic Alzheimer disease diagnosis. Sensors (Basel, Switzerland), 13(5), 6730–6745. https://doi.org/10.3390/s130506730
4. Chen, J. C., Borson, S., & Scanlan, J. M. (2000). Stage-specific prevalence of behavioral symptoms in Alzheimer's disease in a multi-ethnic community sample. American Journal of Geriatric Psychiatry, 8(2), 123–133. https://doi.org/10.1097/00019442-200005000-00007
5. Herrup, K. (2015). The case for rejecting the amyloid cascade hypothesis. Nature Neuroscience, 18(6), 794–799. https://doi.org/10.1038/nn.4017
6. Leissring, M. A., Farris, W., Chang, A. Y., Walsh, D. M., Wu, X., Sun, X., Frosch, M. P., & Selkoe, D. J. (2003). Enhanced proteolysis of β-amyloid in APP transgenic mice prevents plaque formation, secondary pathology, and premature death. Neuron, 40(6), 1087–1093. https://doi.org/10.1016/S0896-6273(03)00787-6
7. Chen, G.-F., Xu, T., Yan, Y., Zhou, Y.-R., Jiang, Y., Melcher, K., & Xu, E. (2017). Amyloid beta: Structure, biology and structure-based therapeutic development. Acta Pharmacologica Sinica, 38. https://doi.org/10.1038/aps.2017.28

DOI: 10.36838/v4i6.10

8. Haass, C., Kaether, C., Thinakaran, G., & Sisodia, S. (2012). Traf ficking and proteolytic processing of APP. Cold Spring Harbor Perspectives in Medicine, 2(5), 1–26. https://doi.org/10.1101/csh perspect.a006270
9. Li, N.-M., Liu, K.-F., Qiu, Y.-J., Zhang, H.-H., Nakanishi, H., & Qing, H. (2019). Mutations of beta-amyloid precursor protein alter the consequence of Alzheimer's disease pathogenesis. Neural Regeneration Research, 14(4), 658–665. https://doi.org/10.4103/1673-5374.247469
10. Sperling, R. A., Aisen, P. S., Beckett, L. A., Bennett, D. A., Craft, S., Fagan, A. M., Iwatsubo, T., Jack, C. R. J., Kaye, J., Montine, T. J., Park, D. C., Reiman, E. M., Rowe, C. C., Siemers, E., Stern, Y., Yaffe, K., Carrillo, M. C., Thies, B., Morrison-Bogorad, M., … Phelps, C. H. (2011). Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's & Dementia: The Journal of the Alzheimer's Association, 7(3), 280–292. https://doi.org/10.1016/j.jalz.2011.03.003
11. Kim, J., Chakrabarty, P., Hanna, A., March, A., Dickson, D. W., Borchelt, D. R., Golde, T., & Janus, C. (2013). Normal cognition in transgenic BRI2-Aβ mice. Molecular Neurodegeneration, 8(1), 15. https://doi.org/10.1186/1750-1326-8-15
12. Bancher, C., Brunner, C., Lassmann, H., Budka, H., Jellinger, K., Wiche, G., Seitelberger, F., Grundke-Iqbal, I., Iqbal, K., & Wisniewski, H. M. (1989). Accumulation of abnormally phosphorylated τ precedes the formation of neurofibrillary tangles in Alzheimer's disease. Brain Research, 477(1–2), 90–99. https://doi.org/10.1016/0006-8993(89)91396-6
13. Brion, J. P. (1990). Molecular pathology of Alzheimer amyloid and neurofibrillary tangles. Seminars in Neuroscience, 2, 89–100.
14. Binder, L. I., Guillozet-Bongaarts, A. L., Garcia-Sierra, F., & Berry, R. W. (2005). Tau, tangles, and Alzheimer's disease. Biochimica et Biophysica Acta - Molecular Basis of Disease, 1739(2), 216–223. https://doi.org/10.1016/j.bbadis.2004.08.014
15. Wiśniewski, H. M., Narang, H. K., & Terry, R. D. (1976). Neurofibrillary tangles of paired helical filaments. Journal of the Neurological Sciences, 27(2), 173–181. https://doi.org/10.1016/0022-510X(76)90059-9
16. Weingarten, M. D., Lockwood, A. H., Hwo, S. Y., & Kirschner, M. W. (1975). A protein factor essential for microtubule assembly. Proceedings of the National Academy of Sciences of the United States of America, 72(5), 1858–1862. https://doi.org/10.1073/pnas.72.5.1858
17. Liu, F., Li, B., Tung, E.-J., Grundke-Iqbal, I., Iqbal, K., & Gong, C.-X. (2007). Site-specific effects of tau phosphorylation on its microtubule assembly activity and self-aggregation. The Europe an Journal of Neuroscience, 26(12), 3429–3436. https://doi.org/10.1111/j.1460-9568.2007.05955.x
18. Iqbal, K., Liu, F., Gong, C.-X., & Grundke-Iqbal, I. (2010). Tau in Alzheimer disease and related tauopathies. Current Alzheimer Research, 7(8), 656–664. https://doi.org/10.2174/156720510793611592
19. Lyman, M., Lloyd, D. G., Ji, X., Vizcaychipi, M. P., & Ma, D. (2014). Neuroinflammation: the role and consequences. Neuroscience Research, 79, 1–12. https://doi.org/10.1016/j.neures.2013.10.004
20. Terry, R. D., Gontas, N. K., & Weiss, M. (1964). ULTRASTRUC TURAL STUDIES IN ALZHEIMER'S PRESENILE DE MENTIA. The American Journal of Pathology, 44(2), 269–297.
21. Carson, M. J., Thrash, J. C., & Walter, B. (2006). The cellular response in neuroinflammation: The role of leukocytes, microglia and astrocytes in neuronal death and survival. Clinical Neuro science Research, 6(5), 237–245. https://doi.org/10.1016/j.cnr.2006.09.004
22. Morales, I., Guzmán-Martínez, L., Cerda-Troncoso, C., Farías, G. A., & Maccioni, R. B. (2014). Neuroinflammation in the pathogenesis of Alzheimer's disease. A rational framework for the search of novel therapeutic approaches. Frontiers in Cellular Neuroscience, 8, 112. https://doi.org/10.3389/fncel.2014.00112
23. Hanisch, U.-K., & Kettenmann, H. (2007). Microglia: active sensor and versatile effector cells in the normal and pathologic brain. Nature Neuroscience, 10(11), 1387–1394. https://doi.org/10.1038/nn1997
24. Liu, B., & Hong, J.-S. (2003). Role of microglia in inflammation-mediated neurodegenerative diseases: mechanisms and strategies for therapeutic intervention. The Journal of Pharmacology and Ex perimental Therapeutics, 304(1), 1–7. https://doi.org/10.1124/jpet.102.035048
25. Perry, V. H., & Teeling, J. (2013). Microglia and macrophages of the central nervous system: the contribution of microglia priming and systemic inflammation to chronic neurodegeneration. Seminars in Immunopathology, 35(5), 601–612. https://doi.org/10.1007/s00281-013-0382-8
26. Cunningham, C., Wilcockson, D. C., Campion, S., Lunnon, K., & Perry, V. H. (2005). Central and systemic endotoxin challenges exacerbate the local inflammatory response and increase neuronal death during chronic neurodegeneration. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 25(40), 9275–9284. https://doi.org/10.1523/JNEUROSCI.2614-05.2005
27. Chen, T. Bin, Lai, Y. H., Ke, T. L., Chen, J. P., Lee, Y. J., Lin, S. Y., Lin, P. C., Wang, P. N., & Cheng, I. H. (2020). Changes in Plasma Amyloid and Tau in a Longitudinal Study of Normal Aging, Mild Cognitive Impairment, and Alzheimer's Disease. Dementia and Geriatric Cognitive Disorders, 48(3–4), 180–195. https://doi.org/10.1159/000505435
28. Jack, C. R. J., & Holtzman, D. M. (2013). Biomarker modeling of Alzheimer's disease. Neuron, 80(6), 1347–1358. https://doi.org/10.1016/j.neuron.2013.12.003
29. Mullins, R. J., Diehl, T. C., Chia, C. W., & Kapogiannis, D. (2017). Insulin resistance as a link between amyloid-beta and tau pathologies in Alzheimer's disease. Frontiers in Aging Neuroscience, 9(MAY), 1–16. https://doi.org/10.3389/fnagi.2017.00118
30. Pickett, E. K., Herrmann, A. G., McQueen, J., Abt, K., Dando,O., Tulloch, J., Jain, P., Dunnett, S., Sohrabi, S., Fjeldstad, M. P., Calkin, W., Murison, L., Jackson, R. J., Tzioras, M., Stevenson, A., d'Orange,M., Hooley, M., Davies, C., Colom-Cadena, M., … Spires-Jones,T. L. (2019). Amyloid Beta and Tau Cooperate to Cause Reversible Behavioral and Transcriptional Deficits in a Model of Alzheimer's Disease. Cell Reports, 29(11), 3592-3604.e5. https://doi.org/10.1016/j.celrep.2019.11.044
31. Jackson, R. J., Rudinskiy, N., Herrmann, A. G., Croft, S., Kim, J. M., Petrova, V., Ramos-Rodriguez, J. J., Pitstick, R., Wegmann, S., Garcia-Alloza, M., Carlson, G. A., Hyman, B. T., & Spires-Jones, T. L. (2016). Human tau increases amyloid β plaque size but not amyloid β-mediated synapse loss in a novel mouse model of Alzheimer's disease. European Journal of Neuroscience, 44(12), 3056–3066. https://doi.org/https://doi.org/10.1111/ejn.13442
32. Koffie, R. M., Hashimoto, T., Tai, H.-C., Kay, K. R., Serrano-Pozo, A., Joyner, D., Hou, S., Kopeikina, K. J., Frosch, M. P., Lee, V. M., Holtzman, D. M., Hyman, B. T., & Spires-Jones, T. L. (2012). Apolipoprotein E4 effects in Alzheimer's disease are mediated by synaptotoxic oligomeric amyloid-β. Brain, 135(7), 2155–2168. https://doi.org/10.1093/brain/aws127
33. Barthélemy, N. R., Li, Y., Joseph-Mathurin, N., Gordon, B. A., Hassenstab, J., Benzinger, T. L. S., Buckles, V., Fagan, A. M., Per rin, R. J., Goate, A. M., Morris, J. C., Karch, C. M., Xiong, C., Allegri, R., Mendez, P. C., Berman, S. B., Ikeuchi, T., Mori, H.,

Shimada, H., … Xu, X. (2020). A soluble phosphorylated tau signa ture links tau, amyloid and the evolution of stages of dominantly inherited Alzheimer's disease. Nature Medicine, 26(3), 398–407. https://doi.org/10.1038/s41591-020-0781-z

34. Itagaki, S., McGeer, P. L., Akiyama, H., Zhu, S., & Selkoe, D. (1989). Relationship of microglia and astrocytes to amyloid depos its of Alzheimer disease. Journal of Neuroimmunology, 24(3), 173–182. https://doi.org/10.1016/0165-5728(89)90115-X

35. Perlmutter, L. S., Scott, S. A., Barrón, E., & Chui, H. C. (1992). MHC class II-positive microglia in human brain: Association with alzheimer lesions. Journal of Neuroscience Research, 33(4), 549–558. https://doi.org/10.1002/jnr.490330407

36. Rogers, J., Luber-Narod, J., Styren, S. D., & Civin, W. H. (1988). Expression of immune system-associated antigens by cells of the human central nervous system: Relationship to the pathology of Alzheimer's disease. Neurobiology of Aging, 9(C), 339–349. https://doi.org/10.1016/S0197-4580(88)80079-4

37. Cai, Z., Hussain, M. D., & Yan, L. J. (2014). Microglia, neuroin flammation, and beta-amyloid protein in Alzheimer's disease. In ternational Journal of Neuroscience, 124(5), 307–321. https://doi. org/10.3109/00207454.2013.833510

38. Dani, M., Wood, M., Mizoguchi, R., Fan, Z., Walker, Z., Morgan, R., Hinz, R., Biju, M., Kuruvilla, T., Brooks, D. J., & Edison, P. (2018). Microglial activation correlates in vivo with both tau and amyloid in Alzheimer's disease. Brain, 141(9), 2740–2754. https:// doi.org/10.1093/brain/awy188

39. Marsh, S. E., Abud, E. M., Lakatos, A., Karimzadeh, A., Yeung, S. T., Davtyan, H., Fote, G. M., Lau, L., Weinger, J. G., Lane, T. E., Inlay, M. A., Poon, W. W., & Blurton-Jones, M. (2016). The adap tive immune system restrains Alzheimer's disease pathogenesis by modulating microglial function. Proceedings of the National Academy of Sciences of the United States of America, 113(9), E1316–E1325. https://doi.org/10.1073/pnas.1525466113

40. Prokop, S., Miller, K. R., Labra, S. R., Pitkin, R. M., Hoxha, K., Narasimhan, S., Changolkar, L., Rosenbloom, A., Lee, V. M. Y., & Trojanowski, J. Q. (2019). Impact of TREM2 risk variants on brain region-specific immune activation and plaque microenvi ronment in Alzheimer's disease patient brain samples. Acta Neu ropathologica, 138(4), 613–630. https://doi.org/10.1007/s00401-019-02048-2

Author

The author, Avni Kabra, is a high schooler in Connecticut interested in neuroscience research. This literature review was written to better understand Alzheimer's Disease and it's causes, which can ultimately assist in finding a cure. She hopes to pursue this further with an interest in mathematical application to scientific problems.

DOI: 10.36838/v4i6.10

IJHSR
ijhighschoolresearch.org
■ RESEARCH ARTICLE

Water Based Space-Time Model Depicting Gravitational Lensing

Ketan Srivastava
Delhi Public School Ghaziabad Vasundhara, Sector-9 Vasundhara, Ghaziabad-201012, Uttar Pradesh, INDIA;
srivastava.ketan@hotmail.com

ABSTRACT: Space is enormous and observing the astronomical phenomena in outer space is not easy. Gravitational lensing is a phenomenon where the bending of space-time deviates the light passing nearby from its original path. Thereby, it is hypothesized, that one may use a water-based model that mimics this extravagant phenomenon. A layer of water on a smooth surface with a hydrophobic region mimics the way the light bends around a black hole. Here, the water layer is equivalent to the space-time consortium and the hydrophobic region is equivalent to the event horizon. Based on the model proposed here, it can be concluded that the size or length of the virtual image is directly proportional to the size of the black hole in the 'not aligned' case and inversely proportional to the 'aligned' case. As the study progressed, it was noticed that the virtual image formed at different positions also followed the pattern of magnetic field lines around a bar magnet. The proposed model makes visualization of the space-time consortium concept easier in that it captures the 'event horizon' and 'ergosphere' while it derives its proof from the phenomenon of gravitational lensing (Figure 1).

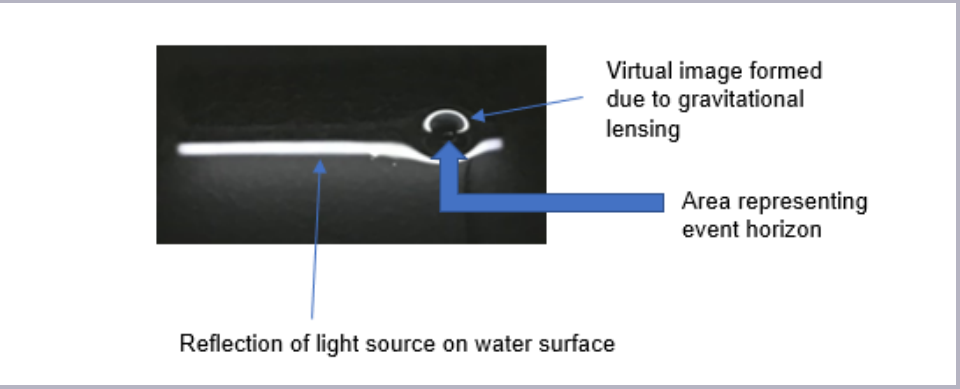


Figure 1: The image of the non-aligned situation of the proposed model is discussed in detail in the results and discussion section.

KEYWORDS: Physics; Astrophysics; Gravitational lensing; Space-time consortium; Event horizon; Ergosphere.

■ Introduction

The famous physicist Albert Einstein helped develop the idea of space-time as part of his theory of relativity.[1] When people talk about space-time, they often describe it as resembling a sheet of rubber. This, too, comes from Einstein, who realized as he developed his theory of general relativity that the force of gravity was due to curves in the fabric of space-time. Objects like the Earth, the Sun, and other massive celestial bodies create distortions in space-time that cause it to bend. These curves, in turn, constrict the ways in which everything in the universe moves, because objects must follow paths along this warped curvature. Motion due to gravity is actually motion along the twists and turns of space-time.[2]

As the light emitted by distant galaxies passes by massive objects in the universe, the gravitational pull from these objects can distort or bend the light (Figure 2). This is called 'gravitational lensing'. Strong gravitational lensing can actually result in such strongly bent light that multiple images of the light-emitting galaxy are formed. Weak gravitational lensing results in galaxies appearing distorted, stretched, or magnified. Although difficult to measure for an individual galaxy, galaxies clustered close together will exhibit similar lensing patterns.[3]

Analyzing the nature of gravitational lensing patterns tells astronomers about the way dark matter is distributed within galaxies and their distance from Earth. This method provides a probe for investigating both the development of structure in the universe and the expansion of the universe. It is apparently impossible to visualize the phenomenon of gravitational lensing or bending of light in the vast universe.[4-7]

DOI: 10.36838/v4i6.11

Figure 2: An image depicting similarity with the proposed model. (A simulated black hole of ten solar masses as seen from a distance of 600 km with the Milky Way in the background). Courtesy: https://theconversation.com/explainer-black-holes-7431.

Method

A thin layer of water on a smooth surface (bathroom tile in this study) with a small dried circular region is created either by dropping a soapy solution (due to surface tension) or sweeping a part of the water layer with a soft hand. The hydrophobic area (HA) thus created is equivalent to the 'event horizon', which when hit by the reflection of the light source mimics the phenomena of gravitational lensing. The circular area around HA is the ergosphere equivalent. Figure 3 represents the cross-sectional view of the model depicting space-time curvature around the black hole and Figure 4 shows the correlation with the proposed model.

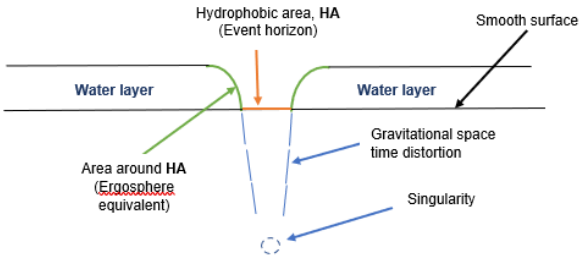


Figure 3: Cross sectional view of the model.

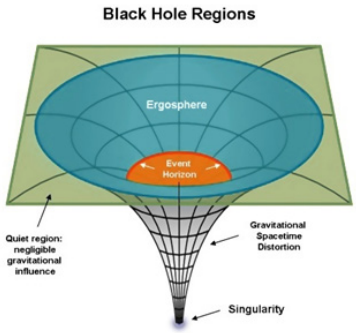


Figure 4: Defining the regions of black hole. Courtesy: https://qph.fs.quoracdn.net/main-qimg-f5e156aa5952733ce726a1f8db60c807-c

The light source used in the experiment is a standard 10W, 240V/50 Hz LED and 54 cm in length and is at a height of 215 cm.

The distance from the base of the wall on which the light source is fixed to the center of the tile on which the experiment is set up is 46 cm (Figure 5). The dimensions mentioned here are not limited to conducting this experiment.

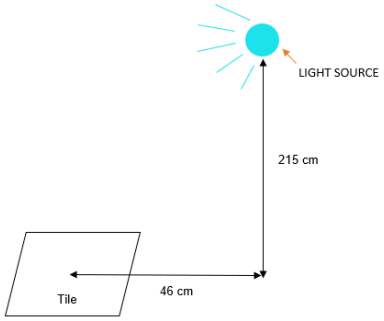


Figure 5: Sketch showing the experimental setup

The observer must position themself in such a way that the circumference of the ergosphere (Figure 4) is just on the tip of the reflected light source and just starts to distort the shape of the reflected light source as in Figure 6a. The length of the virtual image formed opposite to the reflected light source was measured and the observation was recorded. The above-mentioned procedure is good to study the aligned case, however, the observer can position themselves based on the case interested in (not aligned case as in Figure 7a or overlapping the reflected light source case as in Figure 8a).

To prove that the unique effects observed in the case of heavier bodies are different from that of the lighter objects, another experiment was conducted by replacing the hydrophobic area with a lighter object (LO) like a tiny ball of fiber heavy enough to keep a depression on the water layer (Figure 9a). The observer must position themself in the same way as mentioned above to experience the difference.

To experience the resemblance between the virtual image and the magnetic field lines, an experiment was conducted with a colored tip of the reflected light source wherein the bending of the virtual image is following the same pattern as observed in the case of a magnetic field around a bar magnet (Figure 14 a-g).

Results and Discussion

As previously known, due to the presence of massive celestial objects, the light deviates from its path. Einstein proposed the space-time consortium to visualize this effect in which space is imagined as it is placed on a piece of fabric. The curving of space-time depends on the mass of the object.

The different situations this proposed model correlates with are as follows:

A) Virtual image opposite to the reflected light source (in case of heavier celestial objects)

1. Hydrophobic area aligned with the reflected light source
2. Hydrophobic area not aligned with the reflected light source
3. Hydrophobic area overlapping the reflected light source

DOI: 10.36838/v4i6.11

B) Virtual image not opposite to the reflected light source (in case of lighter celestial objects

2. Lighter object (dirt particle in this experiment) not aligned with the reflected light source

3. Lighter object (dirt particle in this experiment) overlapping the reflected light source

A1: Hydrophobic area aligned with the reflected light source:

When the hydrophobic area of the model is aligned with the reflected light source then a virtual image is observed opposite to the reflected light (Figure 6a), also depicted diagrammatically in Figure 6b.

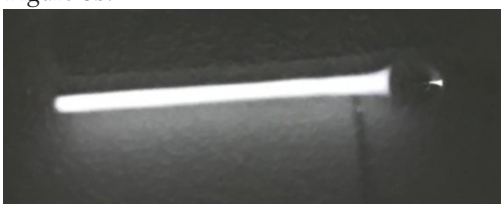


Figure 6a.

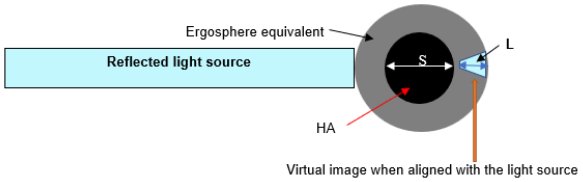


Figure 6b: Diagrammatic representation of Figure 6a.

It is observed that the length of the virtual image (L) is inversely proportional to the size of the HA (S) as mentioned in Table 1.

Length = L

Size = S

L α 1/S

A2: Hydrophobic area not aligned with the reflected light source:

In case the reflected light source is not aligned, then the virtual image is observed as a curved line opposite to the reflected light source (Figures 7a and 7b).

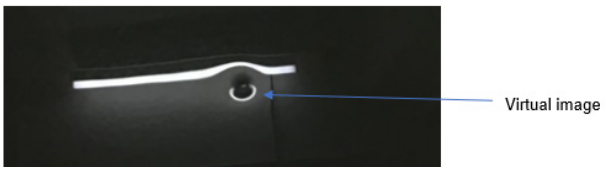


Figure 7a.

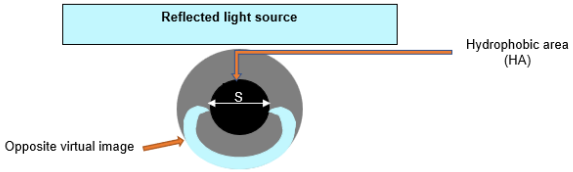


Figure 7b: Diagrammatic representation of Figure 7a.

It is observed that the length of an arc of the virtual image is directly proportional to the size of the hydrophobic area (S).

A3: Hydrophobic area overlapping the reflected light source:

If the HA appears to overlap the reflected light source, then the light around the HA shows a combination of effects arising from both the aligned and not-aligned situations on either side (Figures 8a and 8b).

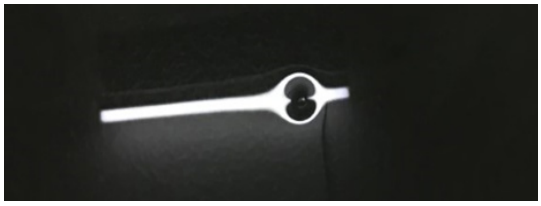


Figure 8a.

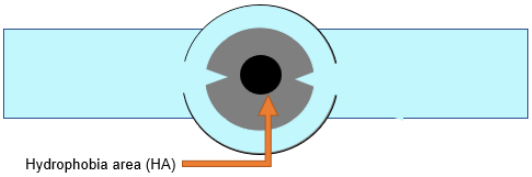


Figure 8b: Diagrammatic representation of Figure 8a.

B1: Lighter object (dirt particle in this experiment) aligned with the reflected light source:

It was observed that the virtual image of the reflected light source is on the same side of the light source (Figures 9a and 9b) in the aligned case.

Figure 9a.

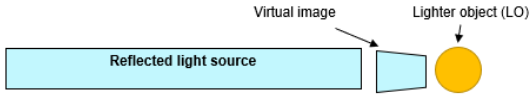


Figure 9b: Diagrammatic representation of Figure 9a.

B2: Lighter object (dirt particle in this experiment) not aligned with the reflected light source:

It was observed that the virtual image of the reflected light source is on the same side of the light source (Figures 10a and 10b) in the not aligned case.

Figure 10a.

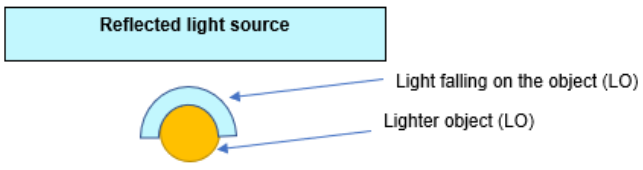


Figure 10b: Diagrammatic representation of Figure 10a.

B3: Lighter object (dirt particle in this experiment) overlapping the reflected light source:

In case the lighter object overlaps the reflected light source based on our view, the object blocks the source of light, and it appears that the light source is broken into two parts (Figures 11a and 11b).

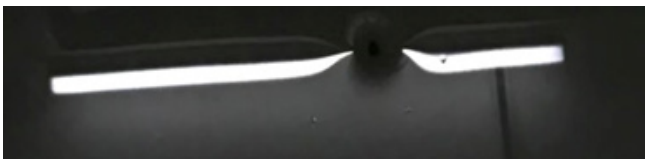


Figure 11a.

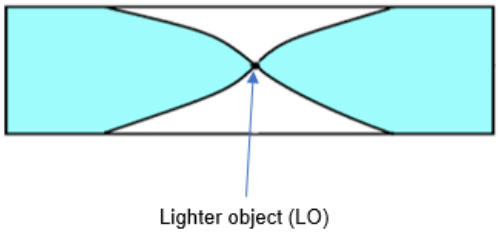


Figure 11b: Diagrammatic representation of Figure 11a.

Virtual image and its resemblance with magnetic field lines:

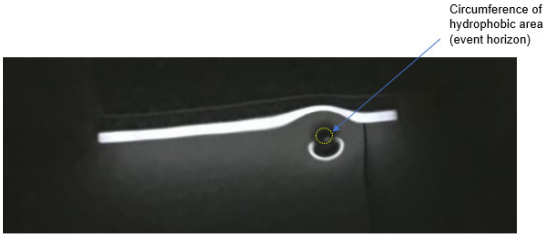


Figure 12.

As observed in Figure 12, the curve of the virtual image resembles the shape of the magnetic field lines around a bar magnet (Figure 13).

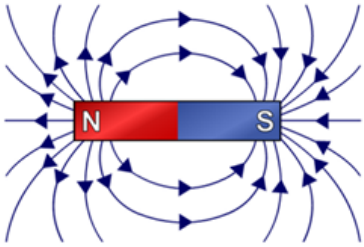
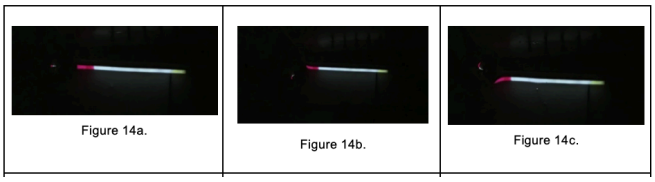


Figure 13.

Figure 14 depicts the pattern observed with the colored tip experiment with different viewing angles and shows the resemblance with the magnetic field lines. .



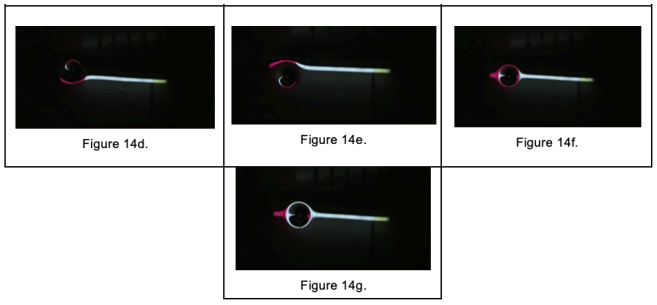


Figure 14 (a-g).

Data Figures:

The data shown in Table 1 and the graphical representation (Figure 15) show the relationship between the size of the HA (event horizon) and the length of the virtual image in case of reflected light in the aligned case. As predicted, the virtual image is inversely proportional to the diameter of HA.

Table 1: Data presenting the Relationship between the size of the HA and the length of the virtual image in case of reflected light in an aligned case.

Entry	Diameter of Hydrophobic Area (mm)	Error for HA diameter (± mm)	Relative error in diameter	Length of virtual image (mm)	Error for virtual image (± mm)	Relative error in virtual image	Relative error of diameter and virtual image	Ratio (image /width)	Error in ratio
1	0.6	0.1	0.17	4.0	0.1	0.03	0.19	6.67	1.28
2	0.8	0.1	0.13	3.9	0.1	0.03	0.15	4.88	0.74
3	1.0	0.1	0.10	4.0	0.1	0.03	0.13	4.00	0.50
4	1.0	0.1	0.10	3.8	0.1	0.03	0.13	3.80	0.48
5	1.0	0.1	0.10	3.7	0.1	0.03	0.13	3.70	0.47
6	2.0	0.1	0.05	3.5	0.1	0.03	0.08	1.75	0.14
7	2.0	0.1	0.05	3.2	0.1	0.03	0.08	1.60	0.13
8	2.1	0.1	0.05	3.9	0.1	0.03	0.07	1.86	0.14
9	2.5	0.1	0.04	3.6	0.1	0.03	0.07	1.44	0.10
10	3.5	0.1	0.03	3.5	0.1	0.03	0.06	1.00	0.06
11	4.5	0.1	0.02	3.1	0.1	0.03	0.05	0.69	0.04
12	4.8	0.1	0.02	3.0	0.1	0.03	0.05	0.63	0.03
13	5.0	0.1	0.02	2.9	0.1	0.03	0.05	0.58	0.03
14	6.0	0.1	0.02	3.0	0.1	0.03	0.05	0.50	0.03
15	7.0	0.1	0.01	2.8	0.1	0.04	0.05	0.40	0.02
16	7.9	0.1	0.01	2.1	0.1	0.05	0.06	0.27	0.02
17	8.9	0.1	0.01	2.6	0.1	0.04	0.05	0.29	0.01
18	10.0	0.1	0.01	2.0	0.1	0.05	0.06	0.20	0.01
19	11.0	0.1	0.01	2.2	0.1	0.05	0.05	0.20	0.01

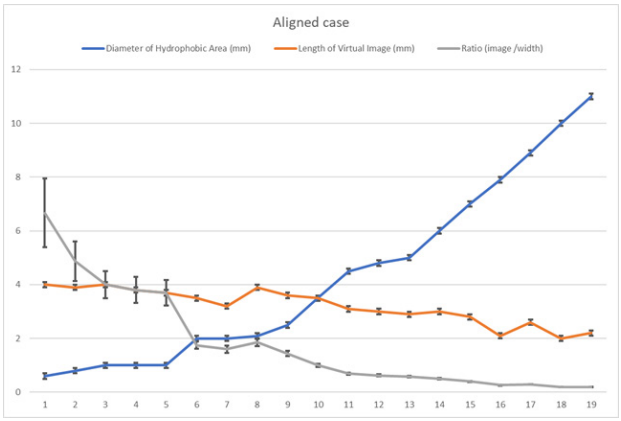


Figure 15: Graphical representation of the relationship of the size of the HA and the length of the virtual image in case of reflected light in aligned case.

Conclusion

The proposed model has a similarity with the space-time fabric model with the upper layer of water mimicking the fabric due to the phenomenon of surface tension. The model can help beginners interested in astrophysics and astronomy to study the concept of 'gravitational lensing'. The observations made in this study are in line with the animated simulation of gravitational lensing caused by a black hole going past a background galaxy.[8] Of the various possibilities, when a hydrophobic area **(HA)** mimicking the event horizon is aligned with the light source then the virtual image is observed op

DOI: 10.36838/v4i6.11

posite to the **HA**. The length of the virtual image formed opposite to the **HA** is inversely proportional to the size of the **HA**. When the **HA** is not aligned with the light source then an arc is observed opposite to it (Figure 7a) and in this case, the length of an arc is directly proportional to the size of the hydrophobic area.

The colored film experiment further shows that the virtual image of light is formed opposite to the light source (Figure 14). In the case of the lighter object **(LO)**, the virtual image of the reflection of light is on the same side of the light source (Figures 9a and Figure 10a). In case the **HA** is overlapping the reflected light source, it shows the combination of effects arising from both the aligned and not aligned cases (Figure 8a), whereas in the case of lighter objects **(LO)**, the reflected light is blocked (Figure 11a). It is also observed that the pattern of virtual images corresponds to the magnetic field lines of the bar magnet and hence hypothesizing it to be having some sort of relationship.

Acknowledgements

I thank my brother Kunal Srivastava for his constant support and for helping me in setting up the experiments. Due credit goes to my entire family for necessary guidance and constant motivation. Special thanks to my mother Hridaya Jyoti Srivastava and father Anil Srivastava who helped me in putting together this manuscript and identifying the gaps during my study towards this experiment.

References

1. Einstein A. (1916), Relativity: The Special and General Theory (Translation 1920), New York: H. Holt and Company.
2. Adam Mann – What is space-time? – Life science contributor https://www.livescience.com/space-time.html
3. https://www.science.org.au/curious/space-time/gravitational-lensing#:~:text=As%20the%20light%20emitted%20by,This%20is%20called%20gravitational%20lensing.
4. Bin Chen, Ronald Kantowski, Xinyu Dai. A Simple Gravitational Lens Model For Cosmic Voids.The Astrophysical Journal, 804, 130, 2015.
5. Markus Selmke - An optical n-body gravitational lens analogy. American Journal of Physics 89, 11, 2021.
6. Massimo Meneghetti, Guido Davoli, Pietro Bergamini, Piero Rosati, Priyamvada Natarajan, Carlo Giocoli, Gabriel B. Caminha, R. Benton Metcalf, Elena Rasia, Stefano Borgani, Francesco Calura, Claudio Grillo, Amata Mercurio, Eros Vanzella. An excess of small-scale gravitational lenses observed in galaxy clusters. Science 11 Sep 2020: Vol. 369, 6509, 1347-1351.
7. Leon Koopmans, Roger Blandford. Gravitational Lenses. Physics Today 57, 6, 45, 2004.
8. https://upload.wikimedia.org/wikipedia/commons/0/03/Black_hole_lensing_web.gif

Author

Ketan Srivastava is presently in twelfth grade at Delhi Public School Ghaziabad Vasundhara in the northern part of India. He completed his class 10 in 2021 and was one of the top scorers in India scoring 499 marks out of 500 (99.80%) in his CBSE exams. He wants to study astrophysics in his higher education and contribute to this field. While he believes he has presented something novel here, there might be more such studies that he might have missed out. With proper guidance, Ketan aims to get a deeper understanding of the topic.

He loves playing and watching soccer and tennis and his favorite pastime is cubing (Rubik's cube).

IJHSR
ijhighschoolresearch.org
■ RESEARCH ARTICLE

Development of Remote Schooling Helper Application

Geonhun Lee
Bergen Catholic High School, 1040 Oradell Ave, Oradell, New Jersey, 07649, USA; geonhun.kevin.lee@gmail.com

ABSTRACT: This paper attempts to solve the issue of teachers' inability to monitor students' attention effectively in the virtual classroom setting which was changed by COVID-19. To solve this issue, this research designs and implements software that can monitor students' postures by using gyroscope sensors and detect students' presence in the class by using webcams. The gyroscope sensor and the webcam send sampled data to the web server where the attention is categorized (as "Out of class," "Present," "Concentrating," "Participating") using the algorithm based on the inclination of the upper body and the location and size of the face detected on the camera. The categorized conjectured attention is shown on the teacher's monitor via a website. After the design and the implementation, this paper runs a mock virtual class with the use of the developed application to verify its effectiveness of the developed application. The results show that an application like this can be used to track students' conjectured attention in a virtual classroom setting and assist teachers in the virtual setting.

KEYWORDS: Systems software; Application; Remote Schooling; Face detection; Posture Monitor; Python.

■ Introduction

A. Background:

COVID-19 has been the single most impactful event in the 21st century. With the new social distancing policies, people had to quarantine themselves in their respective places. Thus, workplaces and schools were forcibly moved to the remote environment in the comfort of people's rooms. Even as the pandemic lengthens and the policy gets more lenient, some people still stay at home and conduct their activities virtually.

And this is the same for the students. According to the study by Burbio, 52 percent of students in the USA planned to do virtual classes, 19 percent are doing hybrid, and only 25 percent plan to do offline lessons in the fall semester of 2020 with the assumption that the situation will be constantly changing.[1] With 2022 approaching, many students are still taking classes online from the comfort of their rooms and the situation will not change for many for some time.

Even though remote activities were inevitable for many students around the nation, they came with their own set of problems, especially for both the students and teachers in the remote class setting. Some problems for students include bad posture and losing focus in the remote class setting while teachers cannot notice and manage students losing focus in the class along with bad posture.

Indeed, one of the problems that students face is bad posture resulting from prolonged sitting. Even before the pandemic, students on average spent 10 hours in a chair.[2] However, with the lockdown caused by COVID-19, students were expected to spend even longer times in chairs. With prolonged sitting in chairs, students are bound to acquire bad posture that negatively affects their health. According to Ma, even if a person sits in a correct position, in the beginning, most people soon revert to a position where the center of mass is shifted to one side of the body within seconds.[2] The position in which the center of mass is shifted to one side can affect one's health negatively by changing the shape of the spine from an S-shape to a C-shape. The C-shape may press the nerves in the back and cause a herniated disc.[2] Not only the back can be affected but also the cervical spine can be affected by the neck pressing it due to VDT syndrome from the prolonged problematic posture.[3] As such, bad posture has a negative influence on health. Yet, this problem is exacerbated by students' prolonged sitting on chairs in a remote class setting.

Poor posture caused by using a computer for a prolonged time promotes the poor performance of the users as well. According to Straker, when the shoulder flexion is greater than 30 degrees, which means that the shoulder is leaning forward 30 degrees, subjects showed poorer performance on visual display unit work. Along with poorer performance, there was greater muscle discomfort for the subjects.[4] Therefore, students will eventually lose their good posture from prolonged sitting at a computer every day and thus perform poorly in their academic careers due to discomfort and muscle fatigue.

In addition, students can lose focus on important tasks of the day due to a phenomenon called "zoom fatigue." Zoom fatigue is an arising phenomenon among the virtual platform users that describes, "the tiredness, worry, or burnout associated with overusing virtual platforms of communication."[5] As many students are overusing virtual platforms every day, many students are feeling burnout from video calling with their teachers, friends, and more. According to Harvard Business Review, virtual platforms are more taxing than physical activities due to a few reasons: more factors of distraction, lack of points of focus, and temptation of multitasking.[6] These factors make remote students put more energy into focusing on remote classes; therefore, remote classes exhaust the students at the end of the day more than face-to-face classes do. With greater fatigue every day, students may lose focus on their academic work. Thus, students can perform even worse in their virtual school than they would have done in the physical school.

DOI: 10.36838/v4i6.12

For people who must use computers for multiple hours a day, these problems are inevitable, yet it is hard to fix since no one can directly notice these problems and tell the person over the screen. Normally, fixing students' posture and guiding them back into focus on the lesson would be the teacher's job. Normally, teachers can usually scan the classroom and immediately find who is not focusing and can guide them back to the lesson of the day. However, this job is much harder now. The average number of students in a class in the US is 24.3.[7] In a virtual class, the teacher must scroll through small windows of students while they are teaching to ensure that the students are participating in the class. In addition, the teacher cannot tell exactly if a student is focusing on the class or just doing something else on the computer. If the teacher cannot monitor students effectively as they do physically, students' postures will eventually become poorer students, and they will lose concentration and perform poorly, academically.[8] Consequently, if teachers were to be able to monitor students better, they would be able to emulate their immediate scanning of students in the live classroom and prevent students from performing poorly.

B. Research Question

Based on the discussion above, the following question can be asked: "Is there a way for a teacher to effectively monitor students' attentions in a virtual classroom setting?"

C. Background Research

Students' attention in virtual classrooms is an understudied topic. Yet, there is a need in allowing teachers to observe the status of all students all at once on a single display to replicate the teacher's scanning of the class.

The student's attention in the class can be measured through blinking eyes, eye movement, brain waves[9], and screen gaze.[10] However, these data alone cannot determine what kind of attention these indicators show.[11] Thus, we believe that measuring sitting posture and analyzing the data to determine whether the student is watching a virtual class will be more useful value for monitoring the student's attention.[11] To develop such a system, the researcher has considered previous literature on how students' attentions and posture can be monitored using camera screens and sensors in the new virtual setting.

One way to monitor student posture in a virtual setting is through a gyroscope sensor. According to the literature, there have been experiments that involve fixing posture using gyroscope sensors, accelerometers, or pressure sensors. According to Hong, gyroscope sensors can be used to determine the neck's inclination for cervical hyperlordosis.[3] According to Ma, an accelerometer, which is very similar to gyroscope sensors, can be put adjacent to cervical vertebrae to determine the change in the vertebrae as the posture changes. There also have been experiments using pressure sensors to log the change in the weight distribution when a person is sitting as posture changes.[2,12]

Since this study needs to monitor students' screen gaze, the degree of bowing the neck should be evaluated. Thus, gyroscope sensors[13], which is a useful tool to measure the inclination of the neck, are optimal to be used.

Another important tool used to monitor students is image processing algorithms. Because a camera is often a requirement in a virtual classroom, facial recognition algorithms[14] can be used to determine if the person is within the camera's sight during the virtual class. Nowadays, there are many facial recognition algorithms available.

Methods

A. System Design:

a. System Overview

The system that I develop includes the following items: MPU6050 gyroscope sensor, Logitech C920 Webcam, Raspberry Pi 4 B, an application developed using Python, and a website and monitoring server based on Python, HTML, and JavaScript. In this paper, this system will only be tested on me at my home during a mock virtual class setting (Figures 1- 2).

Figure 1: The completed image of the system developed.

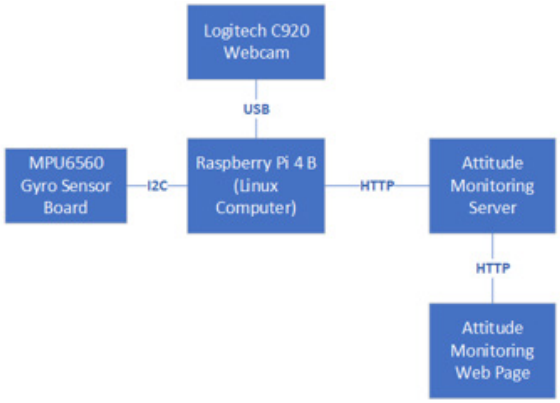


Figure 2: The diagram of the system planned to be developed.

b. Face detection Module Design

The first major component of the system is the facial detection module to identify the student's face during the virtual class. This system uses a separate web camera (Logitech C920 webcam) on the computer since Zoom prioritizes the use of the built-in camera for its use so this system could not access the camera while the Zoom class was functioning. The face detection software was programmed as a Raspberry Pi[15] that receives the image from the camera and computes the coordinates in the camera's FOV of the person's face by using OpenCV Image Processing Library's Face Detection Algorithm.[16]

To appropriately evaluate the student's attention, I have done the following experiments. The experiment consisted of determining the location of the face concerning the camera's field of view, horizontal (Figures 3-4) and vertical angle (Fig-

ures 5-6) that the camera and the face look out on, and the distance of the face (Figure 7) that can be detected from the camera using the face detection algorithm. These experiments set basic standards for the algorithm that determines good behavior

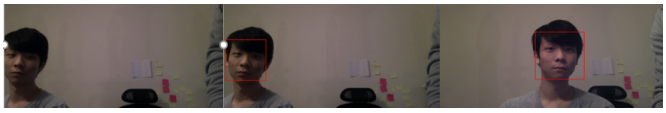


Figure 3: Various Cases of Facial Detection in the horizontal direction at the front-facing position.

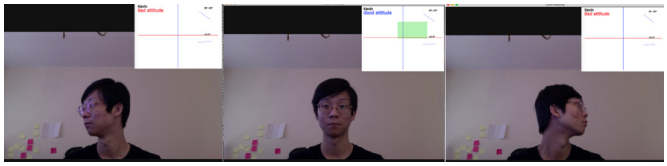


Figure 4: Various Cases of Facial Detection at different horizontal angles of facing positions.

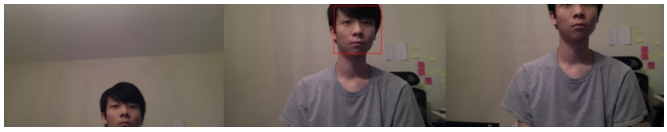


Figure 5: Various Cases of Facial Detection in the vertical direction at the front-facing position.

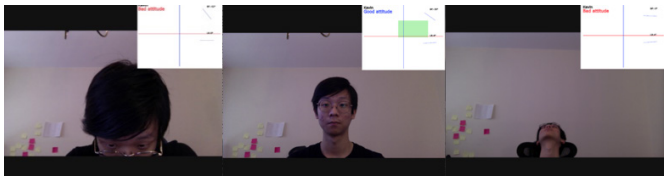


Figure 6: Various Cases of Facial Detection at different vertical angles of facing positions.

Figure 7: Various Cases of Facial Detection at different distances from the camera (50 cm, 1 m, and 2 m from the camera respectively).

c. Posture Monitor Module Design:

The second major component of the system is the 3-axis MPU6050 gyroscope sensor attached to the student's body to compute the orientation of the upper body (Figure 8). It is attached between numbers 5 and 6 of the cervical spines based on previous literature.[3] The BF-axis of the gyroscope faces the student's screen while the IR axis faces perpendicular to the screen. Because the sensor is small, it doesn't impact the student's posture and doesn't give discomfort. Raspberry Pi receives the computed value from the gyroscope sensor.

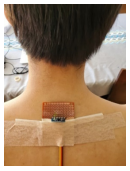


Figure 8: Gyroscope sensor attached to the student's body.

Because there was no data related to healthy posture based on the measurements from the gyroscope, the researcher has conducted experiments to determine the measurements of posture using the gyroscope. For this experiment, the researcher has defined eight postures using three standards based on the gyroscope's measurable range, healthy posture, and facial detection within the field of view of the camera.

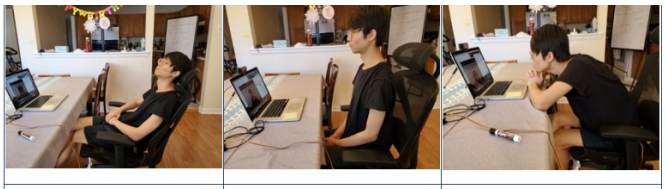


Figure 9: Various backward and forward (BF) inclining postures when the student looks onto the camera.

Figure 9 shows various backward and forwards (BF) inclining postures when the student looks into the camera. The inclination of the posture is measured by the BF-axis of the gyroscope. The range of accepted inclination for the BF-axis is determined by the maximum range in which the face detection program detects the student's face. In posture BF-a, the student faces the front and leans backward. In posture BF-c, the student faces the front and leans forward. In posture BF-b, the student faces the front and does not lean forward at all.

Figure 10: Five different postures when the student is leaning either right or left (LR) while facing front.

Figure 10 shows five different postures when the student is leaning either right or left (LR) while facing the front. These postures are all determined by the inclination on the LR-axis of the gyroscope. The maximum/minimum inclination is determined by the inclination when the student puts his/her head down on either side of the desk, when the student rests his/her chin on their hand, and when the student does not lean at all. Based on these standards, in posture LR-a, the student puts their head down on the right side of the desk. In posture LR-b, the student rests on the chin on the right hand. In posture LR-c, the student does not lead anywhere. In posture LR-d, the student rests their chin on the left hand. In posture LR-e, the student puts their head down on the left side of the desk.

d. Algborithm Implementation

Using the data from the design process of the modules, I have implemented an algorithm to compute the attention of a student based on the posture for a prolonged time.

B. Software Impelementation:

Figure 11 shows the comprehensive configuration of the developed software using the modules and algorithms design-

DOI: 10.36838/v4i6.12

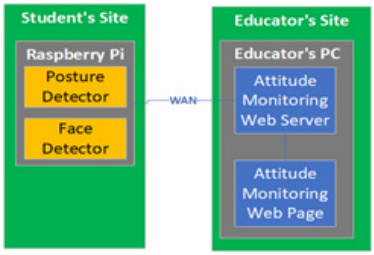


Figure 11: Diagram of software.

ed in earlier steps. On the student's side, Raspberry Pi runs the posture monitor and the face detector. These modules gather images and detect values from the sensors and convert them to use information such as the student's current posture and the face's coordinates in the camera's FOV (Field of view) and send the data out to the Attention Monitoring Server. In the Attention Monitoring Server, the program computes the student's attention using the quantified data on the student's posture and face using its evaluation algorithm and displays the result on its web page. The teacher can directly see the algorithm's evaluation of the student's current attention on the Attention Monitoring Web Page.

C. Application Testing:

The developed application is tested in a mock virtual class setting to ensure that the application properly collects data, computes results, and sends the result to the teacher's webpage. The application is also designed to ensure that various postures are categorized as intended and sends that data to the teacher's webpage. This simulation method verifies the results by comparing the recorded class and the result sent to the teacher.

The simulation consists of a 34-minute virtual Zoom class between the student (researcher) and the teacher (advisor). For the test, the application described in Figure 11 was set up. During the mock class, the application monitors and evaluates students' attention and provides the result to the teacher. All the evaluated attention scores and the mock virtual class are recorded cumulatively. After the mock virtual class, the researcher compared and verified the students' posture recorded on Zoom and the evaluated attention score.

■ Results and Discussion

A. Computation of Reference Value:

The reference value to evaluate students' current attention has been computed using the following method.

a. Face Detection:

The conditions for facial detection are as follows. The camera needs to see the major features of the face and thus the face detection algorithm detects the face within the field of view of the camera. In addition, the detected image size needs to be composed of more than 11.5 percent of the total size of the image since the face detection algorithm can detect the student's face up to 10m from the camera as the student may abuse that feature. The percentage is calculated using the field of view of the camera. Assuming that the student can maximally sit 1m from the camera and the field of view of the camera is 90 degrees, the researcher was able to find the perimeter that the web camera views is π/2m horizontally. Then the researcher calculated the ratio of the perimeter and the average width of a person's face, 0.18m. This results in 0.115, which shows that the face should occupy at least 11.5 percent of the screen width to assume that the student is within 1m from the camera.

b. Range of Inclination of Good Posture:

The following Figure 12 is the result of the experiment to acquire reference values for inclinations of good posture on a vertical axis (axis facing the camera's FOV). In posture BF-a, the furthest the student can lean backward is 45° degrees. In posture BF -b, the student sitting without any leaning on either side produces a value of about -25°~15°. In posture BF -c, the furthest the student can be forward is about -55°.

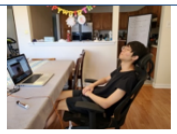
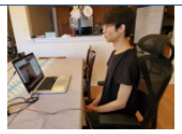
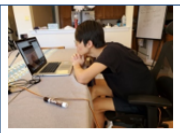
Posture			
Value	45° ~ higher	-25° ~ 15°	-55° ~ lower
type	Posture BF-a	Posture BF-b	Posture BF-c

Figure 12: The result of the experiment to acquire reference values for inclinations of good posture on a vertical axis (axis facing the camera).

The following Figure 13 is the result of the experiment to acquire reference values for the inclination of good posture on the horizontal axis (axis parallel to the camera's FOV). In posture LR-a, the inclination of the student lying on the left side of the table on the arm is lower than -30°. In posture LR-b, the inclination of the student having their chin on the right hand is -30°~-15°. In posture LR-c, the inclination of the student sitting straight is -15°~ 15°. In posture LR-d, the inclination of the student having their chin on the left hand is 15°~ 30°. In posture LR-e, the inclination of the student lying on the right side of the table on their arm is higher than 30°. Based on the result, the reference range of good posture on the vertical axis is set as -30° ~ 30°.

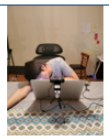
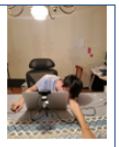
Image of Posture					
Inclination	-30°~ lower	-30°~ -15°	-15°~ 15°	15°~ 30°	30°~ higher
Type	Posture LR-a	Posture LR-b	Posture LR-c	Posture LR-d	Posture LR-e

Figure 13: The result of the experiment to acquire reference values for the inclination of good posture on the horizontal axis (axis parallel to the camera's FOV).

The reference range for both axes is shown in the following chart. The good range of inclination for the vertical axis is -50°~30°and the good range of inclination for the horizontal axis is -30°~30° (Table 1)

Table 1: Reference "Participating" range.

	vertical axis (°)	horizontal axis (°)
Reference "Participating" range	-55 ~ 30	-30 ~ 30

B. Algorithm Based on Reference Value:

Based on the reference values that were computed in the above section, an algorithm has been built to determine a student's current attention status in four levels: "Out of class",

"Present", "Concentrating", and "Participating." The algorithm runs based on a scoring system over a prolonged time. The algorithm takes account of time because a person's posture is constantly changing. For example, a student may drop their pencil or must look behind because someone calls the person. These actions do not necessarily mean that the student is not paying attention but may be deemed as bad conjectured attention consequently. The student starts with a score of thirty at the beginning of monitoring. Then, the sensors sample the student's face location and the angle of the gyroscope. These values are used to compute the student's attention status.

If the posture determined is within the range mentioned in Table 1, the student gets one point. If the student's face is detected and the size of the detection is more than 11.5 percent of the camera's screen, the student gets another point. Every second, the student's facial detection and posture are evaluated and given points from zero to two. The given point is recorded for thirty seconds and is updated every second. The score from the last thirty seconds is used to determine the four levels of student's attention status using the following variables:

CS: Current Status (0~3, 4 levels of status)

SWS: Short-term window size, 5 seconds

LWS: Long-term window size, 30 seconds

SS: Short-term score, 0 ~ 10 points, cumulative score for the last 5 seconds

LS: Long-term score, 0 ~ 60points, cumulative score for the last 30 seconds

SC: Short term criteria, SWS x 2 = 10 points

LC: Long term criteria, LWS + 1 = 31 points

Then, the score is used to determine the four levels of attention as listed in the following:

Level 3: Participating (CS = 3)

Level 2: Concentrating (CS = 2)

Level 1: Present (CS = 1)

Level 0: Out of Class (CS = 0)

The following Figure 14 is the algorithm that uses the variables listed above to compute the student's attention. It was made using Python.

C. Application Testing by Mock Virtual Class Simulation:

```
MINIMUM_FACEWIDTH = 11.5        # minimum face width percent
BF_POSTURERANGE = [-55, 30]     # posture range bf-angle
LR_POSTURERANGE = [-30, 30]     # posture range lr-angle
scoreArr =[1]*30 #: Recent 30 seconds score array
CS = 0          #: Current Score
SWS = 5         #: 5 seconds, Short term window size
LWS = 30        #: 30 seconds, Long term window size
SS = 0          #: 0, Short term Score
LS = 0          #: 0, Long term Score
SC = SWS * 2    #: 5 * 2, Short term Criteria
LC = LWS + 1    #: 30 + 1, Long term Criteria
ATTI = ["Out of Class","Present","Concentrating","Participating"]

while True:
    current = face_detect() + posture_detect()  # 0 or 1 or 2
    scoreArr.push_tail(current)
    scoreArr.pop_front()
    LS=SS=0

    for s in scoreArr:  # long term score accumulation
        LS = LS + s
    for s in scoreArr[-1:(-1 - SWS):-1]:    # short term score acumulation
        SS = SS + s

    if SS == SC:    # perfect during recent 5 seconds
        CS = 3
        # If recent SWS score(SS) is perfect but past 30 seconds record is not good,
        # recover past LWS ~ SWS second data as 1
        if LS < LS:
            for s in range(len(scoreArr) - SWS):
                scoreArr[s] = 1
    elif LS > LC:   # long term score is high
        if SS > SWS:
            CS = 3
        else:
            CS = 2
    elif LS <= LC:  # long term score is low
        if LS >= 5:
            CS = 1
        else:
            CS = 0
    update_ui(ATTI[CS])
```

Figure 14: Pseudocode of the algorithm.

The developed application collected monitoring data for thirty-four minutes of a mock virtual class. The following Figure 15 shows the change in students' attention scores over 34 minutes on the top with the bottom figure zoomed in to show 5 minutes.

Figure 16 shows the image of the student around thirty-one minutes from the beginning of the class (Box B in Figure 15 lower part) which the graph shows as Level 0: "Out of class" attention level. As shown in the image, the student's major facial features were out of the camera's FOV, so the face detection module failed to detect the face and the posture inclination was not within the reference "participating" range (see Table 1).

Figure 15: Change of Attention Score over time for 34 minutes of the mock virtual class.

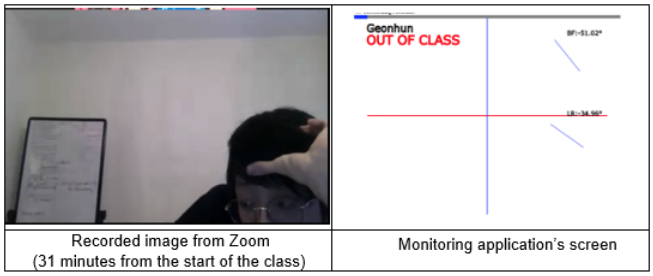


Figure 16: The student's posture and recorded image when the application determined the attention as "Out of Class."

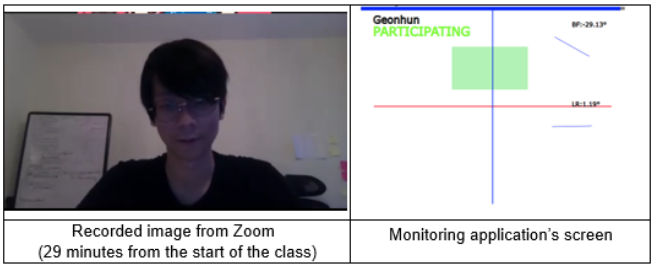


Figure 17: The student's posture and recorded image when the application determined the attention as "Out of Class."

DOI: 10.36838/v4i6.12

Conclusion

In this paper, an application that uses data from a gyroscope sensor module and an image processing module has been developed to assist educators in evaluating students' attention. In addition, a mock virtual class has been conducted to verify that meaningful data can be collected and sent to the educator.

If the developed application is applied in a real virtual class setting, teachers will be able to use less time and effort on monitoring students and thus will be able to focus more on teaching. Moreover, the teacher can give feedback to students based on the developed application so the students can take virtual classes more effectively. In doing so, this application can be proved to be a convenient tool for both students and teachers and promote quality of education in a virtual class setting.

The current research has multiple problems. Therefore, in future studies, we will do the following.

To begin with, a current indication of engagement is not direct or a clear reflection of student engagement. Thus, in the follow-up study, to secure representativeness of the reference value, more trials of the experiment with a greater number of subjects will be performed, deriving the reference value.

Second, although the focus of the study was to promote student attention in the virtual classroom setting, it only succeeded in evaluating student attention and not fully achieving promoting student participation. In future studies, beyond providing one-way information to inform teachers of the students' attention status, the system will be modified and developed so the application can give automatic feedback which will induce students' attention. Also, a two-way information exchange system that enables interaction with teachers will be added to facilitate the teacher's management of students.

Thirdly, it is true that evaluation of a student's attention is only based on face detection and posture detection can be ambiguous. Thus, in future studies, more criteria, such as the movement of eyeballs and perspective, the vibration of heartbeat, and more will be added to make the evaluation of student attention in a virtual classroom even more clear.

Last, of all, the posture monitor module used in this paper did not produce the most accurate evaluation of the student's posture because of a lack of ergonomic research. Therefore, a module that can accurately evaluate one's posture must be developed in future studies.

Acknowledgements

I very much appreciate all those who supported me while researching. My parents, Mrs. Holt, Dr. Esker, Ms. Chung, and Mr. Yoon all supported me to complete this work.

References

1. Liesman, Steve. Half of U.S. Elementary and High School Students Will Study Virtually Only This Fall, Study Shows. *CNBC* [online] Aug 11, **2020**. http://www.cnbc.com/2020/08/11/half-of-us-elementary-and-high-school-students-will-study-virtually-only-this-fall-study-shows.html (accessed Feb 15, 2021).
2. Ma, S.; Hong, S.; Shim, H. M.; Kwon, J.W.; Lee, S. A Study on Sitting Posture Recognition Using Machine Learning. *The Transaction of the Korean Institution of Electrical Engineers*. **2016**, 65(9), 1557-1563.
3. Hong, J.Y.; Lee, M. J.; Yang, H.J; Choi, B.S.; Kim, J. M.; Lee, E.C. A Development of Turtleneck Posture Notification Application Using Gyro Sensor. *Journal of Korea Information Processing Society*, **2018**, 25(2), 860-862.
4. Straker, L.M.; Pollock, C.M.; Mangharam, J.E. The Effect of Shoulder Posture on Performance, Discomfort, and Muscle Fatigue Whilst Working on a Visual Display Unit. I*nternational Journal of Industrial Ergonomics [online]* **1997**, 20(1), 1-10. https://www.sciencedirect.com/science/article/abs/pii/S0169814196000273 (accessed Feb 15, 2021).
5. Lee, J. Neuropsychological Exploration of Zoom Fatigue. *Psychiatric Times [online]* Nov 17, **2020**.
6. Fosslien, L.; Duffy, M.W. How to Combat Zoom Fatigue. *Harvard Business Review [online]* Apr 29, **2020**. https://hbr.org/2020/04/how-to-combat-zoom-fatigue (accessed Feb 15, 2021).
7. Rampell, C. Class Size Around the World. *The New York Times [online]* Sep 11, **2009**. https://economix.blogs.nytimes.com/2009/09/11/class-size-around-the-world/(accessed Feb 15, 2021).
8. Lee, H.; Lee, G.; Kang, S.; Kang, S.; Kwon, M.; Kim, R.; Kim, S.; Kim, S.; Kim, Y.; Jung, D.; Han, E.; Kim, J. Effects of the Home Exercise Program and Exercise Program of Round Shoulder Adjusting on the Shoulder Height, the Level of Trapezius Muscle Activity and Attention Capacity for Middle School Students. J*ournal of The Korean Society of Integrative Medicine*, **2015**, 3(1), 91-103.
9. Ahn, H. Mo.; Nam, S. Ch.; Song ,K. S., Application of bio-signal measurement to identify learning concentration in an e-learning environment. *The Korean Association of Computer Education*, **2012**, 16(2), 125-130.
10. Lee, M.H.; Lee, H.M.; Chung, S.T. Are You Watching Me?: The Design of a Video-based Learning Management System Using L earners' History Data. *Archives of Design Research*, **2021**, 34(4), 225-239.
11. Kim, J. S.; Kim, J.W.; Kim, J.H.; Seo, J.W., EEG & Pitch data-based learning concentration determination system. *Proceedings of the Korean Institute of Information and Communication Sciences Conference*, **2018**, 686-689.
12. Kim, M.; Seo, T.; Lee, J.; Heo, U.; Yoo, H. Development of Smart Sitting Mat using Pressure Sensor for Posture Correction. *Proceedings of the Korean Society of Computer Information Conference*. **2019**, 291-292.
13. Gyro sensors - How they work and what's ahead. Seiko Epson Corp. https://www5.epsondevice.com/en/information/technicalinfo/gyro/(accessed Feb 15, 2021).
14. Golla, R.G. Viola-Jones face detection and tracking explained [Video File] YouTube. Sep 26, 2012. http://www.youtube.com/watch?v=WfdYYNamHZ8(accessed Feb 15, 2021).
15. Raspberry Pi Home Page. http://www.raspberrypi.org/documentation/ (accessed Feb 15, 2021).
16. Face Detection using Haar Cascades. opencv24-python-tutorials. https://opencv24-python- tutorials.readthedocs.io/en/latest/py_tutorials/py_objdetect/py_face_detection/py_face_detection.html (accessed Feb 15, 2021).

Author

The author of the research paper, Geonhun Lee, is currently a student at Bergen Catholic High School. He is interested in Computational biology, the combined application of math, statistics, and computer science to solve problems regarding biology. Researching in this field, he will move forward to achieve his dream.

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Utilizing Doppler Spectroscopy, Transit Photometry, Crowdsourced Home Telescopes, and AI to Find Exoplanets

Logan Reich
Hunter College High School, 71 E 94th St, New York, NY 10128, USA; logantreich@gmail.com

ABSTRACT: Exoplanet detection typically relies on expensive space-based or large ground-based telescopes. A large problem is a lack of telescope time for detection. A network of amateur telescopes, with measuring devices comprised of low-cost spectroscopes and smartphones, was combined with an automated data analysis process involving a custom YOLOv3 AI to collect, clean, and prepare the data. The detection process implemented Doppler spectroscopy and transit photometry, and these two methods were cross-correlated for higher accuracy. In addition, they were used to calculate a variety of orbital parameters and predict exoplanet habitability. This system was then used to evaluate five-star systems, including stars both with and without exoplanets. The AI was 87.63% accurate at identifying stellar spectral lines compared to generally accepted spectra for the stars in question, which shows the high practicality of this system, and a comparison of the calculated values with the generally accepted ones using Spearman's Rho test shows the high accuracy of the system overall, with a p-value of 0.043. Now that this method has been shown to be usable and relatively accurate, this method can be applied to myriad astronomical tasks, such as mapping the spectral classes of star clusters, or scanning for Near-Earth Objects..

KEYWORDS: Astronomy; Astrophysics; Exoplanets; Artificial Intelligence; Crowdsourcing; Doppler Spectroscopy; Transit Photometry.

Introduction

This project attempts to address the fundamental issue in exoplanet discovery: a lack of telescope and observatory time.[1,2] There are an unknown large number of stars that could potentially have exoplanets that have not been investigated for them, and a crowdsourced network would increase the number of telescopes capable of detecting exoplanets. Exoplanet detection is important to provide a better picture of other planets, enable us to understand many aspects of planetary systems, help fight climate change, and aid in the search for extraterrestrial life.[3,4]

While the ability of increased exoplanet detection to improve understanding of planetary systems is immediately clear, its benefits in terms of fighting climate change and aiding the search for extraterrestrial life is less immediately clear. Exoplanet detection can help with climate change by increasing the number of exoplanets available for more detailed follow up studies using precision spectroscopy and photometry to quantify atmospheric dynamics, greenhouse gas concentrations, and temperature profiles in exoplanetary atmospheres.[1,4] Having more varied atmospheric data can provide information on potential feedback loops or climatological interactions that only activate past a certain temperature threshold, allowing for climate modelling to be improved, and maybe even a solution to be discovered.[4] Exoplanet detection can also help with the search for extraterrestrial life by improving targeting techniques as it is believed that a rocky body with liquid water is a prerequisite for the formation of life.[1,3] It could also help find potential alien megastructures such as a Dyson sphere or swarm, suggested to explain abnormal dimming in KIC 8462852 and other irregularly variable stars.[5]

Currently, most detection of exoplanets is done through two approaches: radial velocity and transit photometry.[1,6] Radial velocity techniques involve measuring a star's spectrum over time to detect the shift in wavelength generated by the planet's pull on the star, and calculate the radial velocity and period from this, which allows for many other orbital parameters to be ascertained (with the notable exception of mass, which can only be calculated as the quantity *(mass)* $\sin^3$*(inclination)*.[1] Some recent advances in radial velocity techniques involve precision spectroscopy, allowing for smaller exoplanets to be detected and for exoplanetary atmospheric wind speed to be determined.[7]

Transit photometry involves the measurement of the star's flux over time, and exoplanets are detected through periodic decreases in flux during transits.[1] Some new approaches include "backyard" transit photometry with mirror rigs and CCDs, as well as atmospheric composition measurements.[1,8,9]

This project also suggests stars for further investigation after already having detected a potential indication of an exoplanet's presence and determined approximate orbital parameters, allowing for a much more effective use of professional resources. While there are existing methods allowing people to conduct transit photometry surveys at home, these are often highly complex, and there are not presently any methods of conducting radial velocity surveys at home.[8,9] This project improves upon this by vastly simplifying the process through automation, improving the measurements' accuracy by having more measurements taken from across the network, and by cross-correlating with independent radial velocity measurements. It would also improve upon existing professional methods by reducing time and cost.

DOI: 10.36838/v4i6.13

Methods

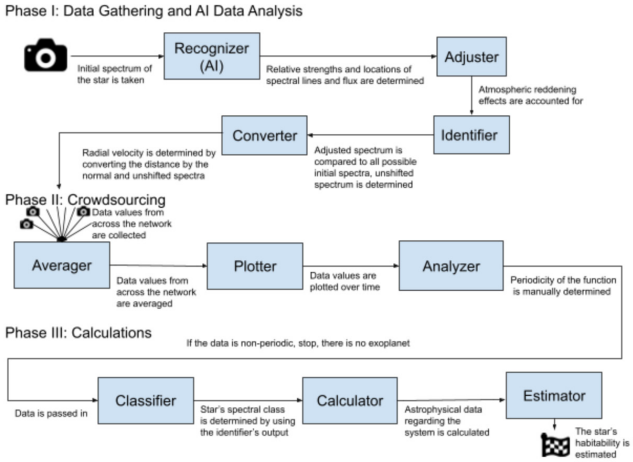


Figure 1: Flowchart showing the overall layout of the methodology.

Figure 2: Image of telescope with the spectrometer attached.

2.1. Data Gathering and AI Data Analysis:

Take the images of the Star's Spectrum

At the beginning of each night the experiment was run, once it was dark, an f/10 Schmidt-Cassegrain 150 mm diameter telescope with a 9 mm Plossl eyepiece was aligned to a star with known apparent magnitude near the star being investigated for calibration, with the entirety of the star's Airy disk visible in the eyepiece, and no other light sources visible. The high-precision visible-light prism spectroscope was mounted to the telescope, and a phone was mounted to the spectroscope with its camera aligned with the spectroscope's eyepiece, as illustrated in Figure 2. A reference photo was taken with the phone camera, recording the spectrum and the intensity of the light. The same process was then used to align the telescope to the star being investigated, with the telescope's auto-tracking feature enabled to keep the telescope pointed at the star under investigation, and a time-lapse phone camera app set up to take a picture of that star's spectrum every 15 seconds with 6-second-long exposures. The spectra were collected by each telescope in the B and V bands and then the spectra in the network were then collected via web upload for AI data analysis (Figure 1).

Figure 3: Image of the interior of the spectrometer, with the sections the AI has recognized as the spectrum for Aldebaran and the reference markers identified in white boxes, as well as a series of calculated values. Note that brightness of the image is significantly enhanced for visibility. This was a calibration photo taken during the study and is meant to demonstrate the appearance of the raw data.

The Recognizer AI Locates the Spectral Lines and Calculates the Flux:

The spectral images of the star being investigated were then loaded into the AI spectrum recognizer. The AI recognized the reference marks, the numbers 4, 5, and 6 and the lines between them printed in the spectroscope and used those to determine the location and strength of the spectral lines by looking at the relative brightness of different areas of the image, as shown in Figure 3. It measured brightness via a direct average of the luminosity of the pixels in the image of the star, adding up the R, G, and B values for each pixel, and averaging over the entire frame. This allowed for an effective calculation of the star's relative luminosity in arbitrary units since all of the light from the star (and no other significant light) is being converted in a predictable manner into the brightness measure. The AI was created in Python using TensorFlow. It was fine-tuned for accuracy and trained on different data sets of images of the spectra of different star classes, and of different stars of the same class, respectively. The spectral recognizer AI used YOLOv3 (You Only Look Once) object detection, which works by using a deep convolutional neural network that splits up an image into grid cells and labels the object it thinks is within that grid cell to predict where the spectrum is, was built with Keras and TensorFlow in Python 3. The YOLO algorithm was useful for this project since it can operate effectively on a small, noisy data set, and does not require a large amount of processing power. The AI was trained on data with a seed from manually analyzed data, which was used to bootstrap to a large enough dataset to train, calibrate, and refine the AI on. All data were gathered from across the network using a Google Form for the images, time and location they were taken, and the star's name. From there, the data were manually verified in order to ensure accuracy. The AI had an accuracy of 87.63%, as measured by the number of correctly predicted data points divided by the total number of data points in a testing dataset. Manual analysis was used to determine what the correct data value was, to the nearest nanometer. All training and testing

datasets used were created by the author using photos taken with the same process but not used in the rest of the study.

The Spectra Are Adjusted for Atmospheric Reddening:

After the AI identified the spectral lines, an adjustment was done for atmospheric reddening by taking the observed magnitude (strength) of the spectral lines and subtracting the first order extinction coefficient multiplied by the airmass. The first order extinction coefficient is 0.4 for the B (Blue) band (400-500 nm) and 0.2 for the V (Visible) band (500-700 nm), while the airmass is the cosecant of the zenith distance, or the angle between the star and the zenith. This angle was calculated as 90 minus the altitude, which was determined via a conversion from the star's right ascension and declination.[10,12] This formula was used as it is a fairly standard way of adjusting for atmospheric extinction that does not require additional measurements, measurement which cannot be taken using at-home equipment.

Identify the Spectral Lines:

Once the spectra were adjusted for atmospheric reddening, the identifier compared the locations and relative strengths of the adjusted spectral lines with all possible combinations of the relative distance between various spectral lines, accounting for all possible effects of line-broadening on the spectrum. After the closest spectrum was determined through a pairwise correlation test, the identifier output the star's predicted unshifted spectrum.

Convert the Shift into Radial Velocity:

The converter then converted the distance between the unshifted spectrum and the shifted spectrum (redshift) into the radial velocity using the standard formula for non-relativistic Doppler shift, $\Delta wavelength/wavelength=c(velocity)$.[1,10,12] In this way, for every telescope in the network, per minute each night, a single radial velocity value and a single relative flux value for the star were determined.

2.2. Crowdsourcing:

Synthesize and Average Data from Across the Network:

At the beginning of each night the experiment was run. After that, the calculated radial velocity and flux values for the star for each telescope in the network were averaged. This resulted in a single radial velocity value and a single flux value for the star per minute each night for the whole network. Had the network been larger, this would have completely normalized the differing random atmospheric effects, since the network should be spaced out sufficiently that, aside for synoptic-scale interactions (which can be considered constant relative to the typically short orbital periods of stars detectable with Doppler spectroscopy or transit photometry), the atmospheric conditions for different observation sites are independent. Due to the independence of the atmospheric phenomena and their inherent randomness, their distribution should mean that they will average out given a large enough sample size.

Plot over Time:

Then averaged radial velocity and flux values were collected over as many nights as possible (in this case seven) and plotted over time. They were graphed as radial velocity vs. time and flux vs. time.

Analyze To See If There Is an Exoplanet:

The graphs were then manually analyzed to see if they formed a periodic function. If so, there was likely to be a massive "Hot Jupiter" (a relatively large gas giant, both in mass and radius, with a short orbital period), exoplanet orbiting the star. By using two entirely independent methods of exoplanet discovery, the presence of an exoplanet was predicted with a much higher confidence. The false positives from smaller secondary stars were caught via the slight increase in relative flux after a transit. If there was an exoplanet, the process proceeded to the calculations.

2.3. Calculation:

Classify the Star's Spectral Type:

Next, by using the output line location references and the unshifted spectrum produced by the identifier AI, the star's unshifted spectrum was compared to references of various stellar spectra (assuming it was a main-sequence star). The spectral class of the reference spectra that the observed spectrum was closest to, based on a pairwise correlation test, was outputted, producing the star's spectral class. The specific references used were a collection of spectra taken as part of this study of five highly visible stars of each spectral class.

Calculate Astrophysical Data for the Star and Exoplanet:

After that, orbital parameters and other data were calculated in the following manner:

Radial velocity semi-amplitude: From the graph of the radial velocity, the radial velocity value that corresponded to the nodes and antinodes of the graph were taken, and the absolute value of the difference of the radial velocity values at the nodes and antinodes (pairing each node with the next occurring antinode) were averaged to get the semi-amplitude.[1,10,12]

Average relative flux: The median of the relative flux values was determined from the relative flux values gathered for the network each minute to get a value to use as the flux for the star.[1,10,12]

Magnitude: The star's apparent magnitude was calculated by comparing the flux measured from the single image of a star with known flux and absolute magnitude taken at the beginning of the night as way to adjust for refraction, and $magnitude_1/magnitude_2=Flux_1/Flux_2$.[1,10,12]

Distance from Earth: The distance from Earth to the star system was determined through the distance modulus, adjusted for interstellar extinction using B-V band estimates based on various surveys of the interstellar medium in the region the star is located.[1,10,12] Absolute magnitude was determined using spectroscopic parallax.[1,10,12] These equations were used as they are standard, precise, and the inputs for them were available from the collected data.

$$magnitude - Magnitude = 5\log\left(\frac{distance}{10\ parsecs}\right) + extinction$$

$$\frac{extinction_v}{Excess_{B-V}} \approx 3$$

Luminosity: The star's luminosity was calculated from its apparent magnitude using a standard equation.[1,10,12]

$$magnitude - magnitude_\odot = -2.5\log\left(\frac{Luminosity}{Luminosity_\odot}\right)$$

DOI: 10.36838/v4i6.13

Mass: The star's mass was calculated using a standard formulation of the empirical luminosity-mass relation for main-sequence stars, as it allowed for a simple conversion and is largely standard.[1,10,12]

$$\frac{Luminosity}{Luminosity_{\odot}} \cong 1.4\left(\frac{Mass}{Mass_{\odot}}\right)^{3.5}$$

Lifetime: The star's lifetime was calculated from a standard formulation of the empirical lifetime-mass relation, as it allowed for a simple conversion and is largely standard. [1,10,12]

$$\frac{Lifetime}{Lifetime_{\odot}} \cong \left(\frac{Mass}{Mass_{\odot}}\right)^{-2.5}$$

Temperature: The star's temperature was estimated from its spectral class via standard values for the temperatures of various spectral classes with O being 30,000 K to 60,000 K, B being 10,000 K to 30,000 K, A being 7,500 K to 10,000 K, F being 6,000 K to 7,500 K, G being 5,000 K to 6,000 K, K being 3,500 K to 5,000 K, and M being less than 3,500 K. [1,10,12]

Radius: The stellar radius was calculated using the Stefan-Boltzmann law.[1,10,12]

$$Luminosity = 4\pi(radius)^2\sigma \cdot Temperature^4$$

Exoplanet radius: The exoplanet radius was calculated using a standard equation, $F_{transit}/F_{average}=R_{planet}/R_{star}$.[1,9,11]

Orbital period: The exoplanet's orbital period was determined as the average of the periods of the radial velocity and relative flux graphs as a better measure of the actual period.[1,10,12]

Semi-major axis: The exoplanet's semi-major axis was determined from Newton's form of Kepler's third law, a standard equation in orbital mechanics.[1,10,12]

$$Period^2 = \frac{4\pi^2 \cdot (semi-major\ axis)^3}{G \cdot Mass}$$

Exoplanet mass: A minimum on the planet's mass was determined by using the star's mass, the orbital period, and the semi-major amplitude of the star's radial velocity, with a standard equation.[1,10,12] Inclination assumed from the fact that it is transiting..[1,9,11]

$$\frac{Mass_{planet}^3 \sin^3(inclination)}{(Mass_{planet} + Mass_{star})^2} = \frac{velocity^3 \cdot period}{2\pi G}$$

Exoplanet orbital velocity: The exoplanet's minimum orbital velocity was calculated from the vis-viva equation, a standard equation in orbital mechanics.[1,10,12]

$$velocity = \sqrt{G(mass)\left(\frac{2}{radius} - \frac{1}{axis}\right)}$$

Exoplanet density: The exoplanet's minimum density was calculated using the definition mass/volume.[1,10,12]

Average exoplanet surface temperature: Using the semi-major axis of the exoplanet's orbit and the star's luminosity, as well as an estimated value for the bond albedo of a "Hot Jupiter" exoplanet (0.75), the exoplanet's average surface temperature was calculated using a standard equation.[1,10,12]

$$Temperature = \sqrt[4]{\frac{Luminosity_{star}(1-albedo)}{16\sigma\pi(radius)^2}}$$

Estimate If the Exoplanet Is Habitable:

Finally, the astrophysical data from above was inputted into the commercially available Universe Sandbox 2 simulation software, to allow for an estimate of the planet's habitability using the inbuilt Goldilocks Zone calculation.

Ultimately, due to limitations from COVID and time constraints, the network ended up consisting solely of observations made by the author, though the network principle of the project was still demonstrated effectively as multiple sets of observations were made and then averaged by taking more frequent photos to effectively generate multiple sets of data.

Results and Discussion

Five-star systems — Pegasi 51, Altair, Deneb, Algol, and Vega — were chosen to test the procedure. These systems were chosen for their relative brightness, and to ensure that some of the stars would have exoplanets in order to ensure that the detection system works, as well as that some stars would not have exoplanets to check for false positives. These five tests were important as practical proof that the approach works in practice.

For each star system, the methodology listed in section 2 was followed exactly, with reference images and spectra being taken, and then data were reduced to convert the input spectra into relative flux and radial velocity values. Those values were then plotted over time and analyzed in order to assess the presence of an exoplanet and calculate orbital parameters and estimate habitability if there was an exoplanet present.

Figures 4, 5, and 6 display examples of the graphs of radial velocity and flux generated for analysis of the exoplanet for three selected star systems. Figure 4 shows the periodic appearance of the radial velocity curve in blue as well as a periodic trend in the relative flux that are in sync, while Figures 5 and 6 display the lack of a periodic trend indicating the absence of an exoplanet. None of the graphs have error bars due to the unknown variability of the measurement system, primarily because of some unknown components in the phone camera used to record the spectra, which is likely to be the limiting factor in the setup. The hypothesis was that the measured and calculated data would be similar to the internationally accepted values within a margin of error, which was validated.

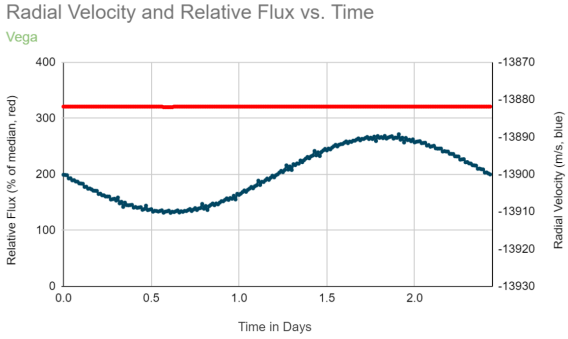


Figure 4: Graph of the Doppler-calculated radial velocity and the adjusted flux of Vega over time. Values are a moving average over 15 minutes. Certain values are interpolated based on other data, since because of daylight, not all of the data was collected at once. The red is the relative flux, and the blue is the radial velocity. This displays the presence of an exoplanet, as indicated by the periodic appearance of the graph.

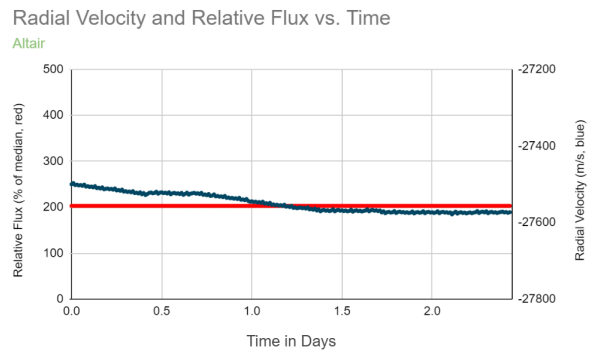


Figure 5: Graph of the Doppler-calculated radial velocity and the adjusted flux of Altair over time. Values are a moving average over 15 minutes. Certain values are interpolated based on other data, since because of daylight not all of the data was collected at once. The red is the relative flux, and the blue is the radial velocity. This displays the lack of an exoplanet, as indicated by the lack of periodicity.

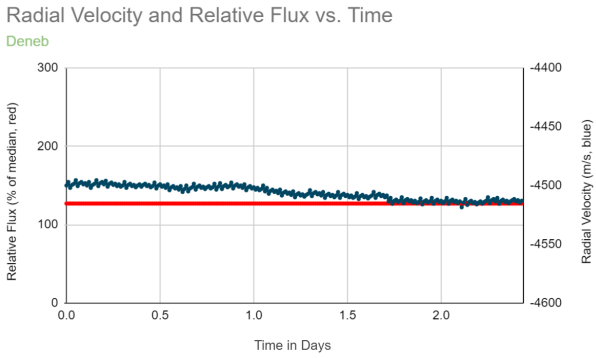


Figure 6: Graph of the Doppler-calculated radial velocity and the adjusted flux of Deneb over time. Values are a moving average over 15 minutes. Certain values are interpolated based on other data, since because of daylight not all of the data was collected at once. The red is the relative flux, and the blue is the radial velocity. This displays the lack of an exoplanet, as indicated by the lack of periodicity.

Table 1 below describes a condensed form of the raw data as well as the calculated parameters for all five-star systems. The calculated and measured values are very close to the internationally accepted ones, with a statistically significant correlation between the accepted and measured values using Spearman's Rho test ($p=0.043<0.05$). This statistically significant correlation shows the validity of this system and calculations as a tool for identifying exoplanet candidates.

Table 1: Collection of key calculated data, as well as an estimate of exoplanet habitability.

Star Names ->	Pegasi 51	Altair	Deneb	Algol	Vega
Peak Wavelength	520 nm*	402 nm	340 nm	224 nm	300 nm*
Radial Velocity Semi-Amplitude	140 km/s*	N/A	N/A	N/A	10 m/s*
Average Relative Flux	44*	203	127	3	321*
Apparent Magnitude	5.5*	0.76*	1.25	2.12	0*
Absolute Magnitude	4.5*	2.22*	-8.38	0	0.5*
Distance From Earth	15 pc*	5 pc	840 pc	30 pc	8 pc*
Luminosity	1.5 L⊙*	10.5 L⊙*	200000L⊙	180 L⊙	40 L⊙*
Mass	1 M⊙*	1.5 M⊙*	20 M⊙	3 M⊙	2 M⊙*
Lifetime	1e10 yrs*	1e8 yrs	1e8 yrs	1e9 yrs	1e9 yrs*
Temperature	5500 K*	7200 K*	8500 K	13000 K	9,600 K*
Radius	1.2 R⊙*	2 R⊙	200 R⊙	3.5 R⊙	3 R⊙*
Exoplanet Radius	1.0 R_j	N/A	N/A	N/A	3 R_j
Orbital Period	4.25 days*	N/A	N/A	N/A	2.43 days
Semi-Major Axis	0.06 AU	N/A	N/A	N/A	0.05 AU
Minimum Exoplanet Mass	0.5 M_j	N/A	N/A	N/A	12 M_j
Minimum Exoplanet Density	0.07 g/cm^3	N/A	N/A	N/A	0.14 g/cm^3
Average Exoplanet Surface Temperature	146 K	N/A	N/A	N/A	419 K
Habitability	Too cold	N/A	N/A	N/A	Too hot

Note (*): Certain outlier data points have been removed, and others interpolated. Stellar data values were also used as a seed to bootstrap data for the AI training in order to calibrate it, as well as used during the manual analysis to generate coeffecients for the comparison equations described in the methods section.

Figures 7 and 8 display the result of modelling the calculated values using Universe Sandbox 2 and predicting the habitable zones around the star. The definition for habitability used here is the presence of liquid water on the surface, a common standard for habitability. It is important to note that this model uses reference albedos for different types of planets and does not account for the greenhouse effect, feedback loops, internal planetary heating, tidal heating, or other possible factors that could generate liquid water or affect the temperature on an exoplanet besides direct absorption of sunlight.

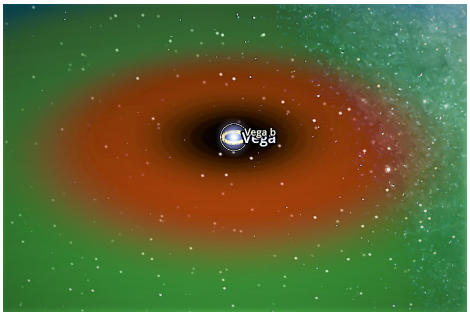


Figure 7: Depiction of the Vega star system with the values calculated for that particular system, showing the habitable zone for that star in green. Red is the slightly too warm zone, and blue is the slightly too cold zone, modeled using Universe Sandbox 2. It shows the one Hot Jupiter exoplanet orbiting Vega, Vega b, and that it is likely too hot to be habitable (Universe Sandbox 2 is a commercially available software that can be used to relatively accurately model astrophysical phenomenon, including star systems).

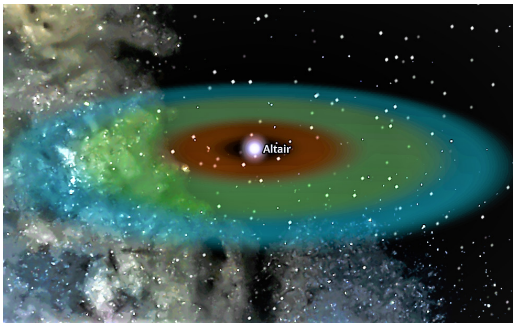


Figure 8: Depiction of the Altair star system with the values calculated for that particular system, showing the habitable zone in green for that star. Red is the slightly too warm zone, and blue is the slightly too cold zone, modeled using Universe Sandbox 2. It depicts the lack of exoplanets (Universe Sandbox 2 is a commercially available software that can be used to relatively accurately model astrophysical phenomenon, including star systems).

Conclusion

This project has shown that amateur home telescopes rigged with inexpensive spectroscopes and phone-photometers can be utilized to discover exoplanets and calculate key factors affecting them by leveraging two popular methods of

DOI: 10.36838/v4i6.13

exoplanet discovery, Doppler spectroscopy and transit photometry. Clearly, as the AI in this system worked to within 87.63% accuracy on average for astrophysical data for five-star systems with only one telescope, it is highly practical and promising. Additionally, comparison of the similarity of the measured values and the internationally accepted ones using Spearman's Rho test revealed that there is a statistically significant correlation between the two sets of values ($p=0.043$). However, it is important to note that these results may be significantly biased due to the extremely limited number of observations and the high margin of error involved. The bias is likely to be introduced from the fact that the actual network did not have the desired geographic diversity. There were three key phases to this project, each with their own benefits: first, gathering of initial data and AI data analysis to generate precise, accurate data at a low cost and with amateur equipment; second, crowdsourcing data to average out random errors; third, calculation of a wide variety of useful data points from a few initial ones.

Now that this method of combining AI, crowdsourcing, and low-cost home equipment has been shown to be effective, reliable, cheap, and relatively easy, it can be applied to a wide variety of other applications in science. The same principles used in this project can be applied to home monitoring systems to track pollution; phone cameras to track things such as crop yields, endangered animals, and water sources; home thermometers to track global warming; and more. Data gathered by this method could be used to provide a more accurate and complete picture of a large number of problems scientists are trying to solve today and help lead to solutions.

At its current stage, this project has many limitations, including the small number of nodes in the network, a lack of geographic diversity in the network, the limited precision of the measurement devices, and the relatively small test size. These limitations could be overcome by a recruitment drive to encourage more people to join the network, which would improve the size of the network, as well as the geographic diversity. A redesign of the spectroscopes or the addition of a secondary lens in front of the phone camera used could help improve the precision of the measurement devices. Lastly, the small test size could be improved simply by investigating more star systems, and a wider variety of star systems. More observations would also allow for more data to be available to train the AI, which could increase its accuracy as well.

While overall this project was successful, it could be improved further by improving the AI to remove large-scale differences more effectively. Some of these large-scale differences include the change in atmospheric reddening at different latitudes, synoptic-scale atmospheric interactions, and systemic errors in the low-cost spectrometers used to generate data that is not only relatively accurate (which was enough to generate useful results for finding exoplanets in this project), but absolutely accurate. Combining this existing system with more precise refraction information from Doppler radar might also be useful in improving accuracy if the network cannot be expanded to a large enough size. Finally, another improvement might be open sourcing the tools used in this project to allow anyone to help discover exoplanets in their backyard. The author is currently working on improving the efficiency, reliability, and documentation of the code until it is at a state where it can be readily accessed online either via a website or public GitHub repository.

In addition, while the purview of this project falls more under the realm of pure science, the astrophysical data of this project and the average surface temperature estimates, combined with an investigation of the atmospheric spectra of the exoplanets using larger telescopes to measure the atmospheric temperature and composition, could be used to measure the effects of greenhouse gases and climate change on other planets. A better understanding of how the greenhouse effect works on other planets could help elucidate feedback loops that could kick in on earth once the temperature rises past a certain point or the levels of certain greenhouse gases pass a threshold. In turn, that could help improve the understanding of climate change and allow for the development of better climate models, and maybe one day allow for the solution to climate change altogether.

Acknowledgements

The author would like to thank Arlene, Evan, and Tina Reich for the telescope used, the smartphone camera, as well as transport to and from the site.

References

1. Dotson Renée, & Seager, S. (2011). Exoplanets. University of Arizona Press.
2. Howell, S. B. (2020). The grand challenges of exoplanets. Frontiers in Astronomy and Space Sciences, 7. https://doi.org/10.3389/fspas.2020.00010
3. Brennan, P. (2019, November 19). Exoplanet exploration: Planets beyond our solar system. NASA. Retrieved December 12, 2021, from https://exoplanets.nasa.gov/news/1610/why-do-scientists-search-for-exoplanets-here-are-7-reasons.amp.
4. Shields, A. L. (2019). The climates of other worlds: A review of the emerging field of exoplanet climatology∗. The Astrophysical Journal Supplement Series, 243(2), 30. https://doi.org/10.3847/1538-4365/ab2fe7
5. Andersen, Ross. "The Most Mysterious Star in Our Galaxy." The Atlantic. Atlantic Media Company, November 30, 2015. https://www.theatlantic.com/science/archive/2015/10/the-most-interesting-star-in-our-galaxy/410023/.
6. Caltech. (n.d.). Exoplanet and candidate statistics. Exoplanet and Candidate Statistics. Retrieved December 12, 2021, from https://exoplanetarchive.ipac.caltech.edu/docs/counts_detail.html.
7. Birkby, J. L. (2018). Exoplanet Atmospheres at High Spectral Resolution. Astro-Ph.EP.
8. Licchelli, D. (2007). Exoplanet photometry from backyard observa tory. Memorie Della Societa Astronomica Italiana Supplementi, 11(211).
9. Munkacsy, M. (2018, December 26). Backyard exoplanet transit observing #1. Astronomical Society of Southern New England. Retrieved December 12, 2021, from https://assne.org/backyard-exoplanet-transit-observing-1/.
10. Karttunen, H. (2003). Fundamental Astronomy. Springer.
11. Castellano, T., & Laughlin, G. (2002). The Discovery of Extrasolar Planets by Backyard Astronomers. The Society for Astronomical Sciences 21st Annual Symposium on Telescope Science, 9–17.
12. Carroll, B. W., & Ostlie, D. A. (2018). An Introduction to Modern Astrophysics. Cambridge University Press.

Author

Logan Reich is a junior at Hunter College High School. He is captain of his school's Science Bowl team, which he captained to top 24 in the nation, and captain of his school's Science Olympiad team, where he placed first in astronomy in New York City North. He also placed in the top 20 in the nation in the National Astronomy Competition, and attended the US Astronomy and Astrophysics Olympiad Summer Training Camp.

DOI: 10.36838/v4i6.13

IJHSR
ijhighschoolresearch.org
REVIEW ARTICLE

Tumor Proliferation Through Sialic Acid Dynamics and Anti-Siglec-Sialoglycans Preventive Strategies

Phuc (James) Nguyen
La Quinta High School, 10372 W McFadden Ave, Westminster, California 92683, USA; jamesnguyen9105@gmail.com

ABSTRACT: Glycans, carbohydrates on cell membranes, are often overexpressed in cancer cells. Sialic acid, a sugar found at the terminal branch of glycans, plays a vital role in carbohydrate-protein interactions, intercellular communication, and bacterial or viral infections. However, hypersialylation of glycans has been linked to hallmarks of cancer such as tumor growth, angiogenesis, metastasis, and resistance to immune cells. Upregulation of sialylation promotes evasion of immune surveillance and survival of malignant cells. Given this role of sialic acid, inhibiting the sialylation of glycans could provide quantum leaps in treating cancer. Over-sialylation is significant to cancer mechanisms (growth, metastasis, and immune evasion); therefore, a high therapeutic value could be found in preventing sialic acid's role in cancer. This paper reviews how sialic acid plays a key role in assisting the growth and survival of cancer cells, despite human immune cells' attacks; it also summarizes findings that encourage removing hypersialiated glycans as a strategy against multiple cancerous cell types.

KEYWORDS: Biomedical and Health Sciences; Genetic and Molecular Biology of Disease; Cancer Progression; Resistance; Sialic Acid; Siglec.

Introduction

Upon every cell's surface are carbohydrates called glycans, many of which can form a layer called a glycocalyx.[1] Glycoprotein, glycolipids, and proteoglycans are all different forms of glycans, which consists of a single unit that makes cellular communication possible.[2]

Sialic acids are attached to the terminal end of glycans through various glycosidic linkages such as α2,3, α2,6, α2,8, and α2,9.[3]

Approximately, fifty derivatives of neuraminic acids can be found under the category of sialic acid.[4] N-acetylneuraminic acid (Neu5Ac) and N-glycolylneuraminic acid (Neu5Gc) is among the most widely-known sialic acid derivatives in mammals.[4] Whereas Neu5Ac has an acetyl group, Neu5Gc has a glycolyl group.[4] Both of these groups are placed on the fifth carbon atom (C5).[4] Since human DNA does not contain a functional cytidine monophosphate N-acetylneuraminic acid hydroxylase (CMAH) gene, which would encode the enzyme forming CMP-Neu5Gc from CMP-Neu5Ac, our cells lack the derivative Neu5Gc.4 Fortunately, Neu5Gc still exists in the cellular membrane's glycome due to daily nutrition intake.[4]

Cancer cells have been discovered to be coated with hyper-glycosylated glycans, and even the tiniest modification to these glycomes can result in a significant transformation from normal to cancerous cells.[5] Aberrant sialylation, truncated O-glycans, and fucosylation is examples of such changes that could lead to the development of cancer tumors.[5] Over-glycosylation is correlated with all of the cancer hallmarks, from cancer cell growth to metastasis to immune evasion to angiogenesis.[6] An upregulation of tumor sialoglycans like SLA, SLS, STn, and GM2 are significant promoters in the advancement of cancer.[7]

Discussion

In the human body, the synthesis of sialic acid begins at UDP-GlcNAc in the cytosol after the glucose has penetrated.[8] The enzyme UDP-GlcNAc 2-epimerase/ManNAc-6-kinase (GNE) is the first to catalyze UDP-GlcNAc, transforming UDP-GlcNAc to ManNAc-6-P.[9] Then, through the help of enzymes Neu5Ac 9-phosphate synthase (NANS) and Neu5Ac-9-phosphate phosphatase (NANP), ManNAc-6-P turns into Neu5Ac.[10] Lastly, Neu5Ac is transported to the nucleus, where it interacts with the cytosine 5′-monophosphate N-acetylneuraminic acid synthetase (CMAS) to construct CMP-Neu5Ac.[11]

From here, the transportation of CMP-Neu5Ac to the Golgi apparatus's glycoconjugates occurs through the help of SLC35A1 sialyltransferases.[11] Through this process, approximately 20 different sialyltransferases, including ST3GAL1-5, ST6GAL1, 2 and ST6GALNAC1-6, and ST8SIA1-6, are supplied with CMP-Neu5Ac as a substrate.[4] The ST8SIA1-6 sialyltransferases generate α2,3-, α2,6-, α2,8, or α2,9-linked sialic acids through connecting Neu5Ac's second carbon to its third carbon, sixth carbon to eighth carbon.[4] Lastly, sialylated glycoconjugates, packaged in vesicles, are secreted out of the membrane.[4]

Alternatively, sialidases (also called neuraminidases) can release sialic acid glycoconjugates.[12] There are four distinct types of sialidases: NEU1, NEU2, NEU3, and NEU4.[13] The sialidase NEU1, which is lysosomal, enables the breakdown of sialoglyco conjugates.[14] The sialidase NEU2, which is cytosolic, prevents interactions with gangliosides.[15] The sialidase NEU3, often found on the plasma membrane, works specifically with gangliosides.[16] The NEU2 and NEU3 function quite differently. And, lastly, the sialidase NEU4, often found in the outer mitochondrial membranes or the lysosomal lumen, is an enzyme whose substrate varies widely; most gly-

DOI: 10.36838/v4i6.14

coproteins, gangliosides, and oligosaccharides can be processed by sialidase NEU4.[17]

After exocytosis outside the cell, the sialic acids are filtered back into the cytoplasm, where they restart their cycle to produce more sialyl glycoconjugates.[18] Otherwise, they would be degraded by N-acetylneuraminate lyase (NAL) to ManNAc and then eventually pyruvate.[18] Since there is a lack of protein transporter on the membrane that specifically works for Neu5Ac or ManNAc, Neu5Ac or ManNAc penetrate the cell mainly through pinocytic or endocytic pathways.[19]

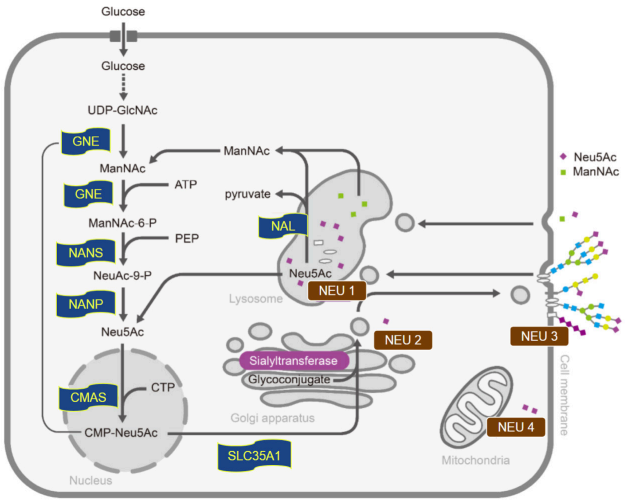


Figure 1: The Life Cycle of Sialic Acid. Zhou, Xiaoman, *et al.* "Biological Functions and Analytical Strategies of Sialic Acids in Tumor." Cells vol. 9, 2 273. 22 Jan. 2020, doi:10.3390/cells9020273

Sialic Acid in Tumor Growth:

Overproduction of sialic acids is a distinct characteristic of cancerous cells during tumor formation and malignant advancement.[20] Hyper sialylation is more seen in tumor tissues than in any other healthy functioning tissue.[20] The total amount of sialic acid present in serum or glycolipid-bound membranes is noticeably high in leukemia, ovarian, breast, colorectal, and pancreatic cancers.[21] In more deadly cancer such as glioma, neuroblastoma, and lung cancer, polysialic acid is overexpressed.[22] Hypersialylation plays major assistance in the acceleration of cancer progression, which eventually results in a poor prognosis.[23] The rise of sialic acids in tumors is a direct consequence of a metabolic flux of sialyltransferases/sialidases.[4] Tumor cells initiate their hypersialytation by increasing their uptake of glucose, the first raw material entering the cell through endocytosis.[4]

Inactivation of the CMAS gene, which codes for a key enzyme that finalizes the production of sialic acid in the nucleus, has been found to inhibit the production and activation of sialic acid.[23] Sialic acid encourages cancer progression through the promotion of tumor expansion, evasion of apoptosis and immune attack, and development of metastases through bloodstream extravasation.[24] Unsurprisingly, sialylation (growth of sialic acid as a terminal sugar on glycans) of proteins can trigger multiple signaling molecules and pathways to further cancer advancement.[24]

For instance, when α2-3-linked glycan is hypersialylated in gastric carcinoma, the activation of receptor tyrosine kinases MET and RON occurs, which results in a more invasive phenotype.[25] Accordingly, N-glycans, when terminally sialylated, can help tumor cells resist hypoxia which confers malignancy to the cell phenotypes.[26] Specifically, the α2-6 sialic acid growing on N-glycans causes the upregulation of E-cadherin, which then promotes the impairment of cancer cell adhesion and aggrandizer metastasis.[27]

Drugs targeting sialic acids are developing to further the prevention of cancer metastasis.[28] *In vivo*, a sialic acid called glycomimetic (P-3Fax-Neu5Ac) has been defined as the promoter of tumor growth.[28] So, delivery of anti-sialic-acid nanoparticles has been performed to stop the metastasis of melanomas to lung cancer in a murine model.[29] Drugs such as Ac53FaxNeu5Ac, when intratumorally injected, can thwart sialic acid expression and therefore suppress tumor growth in many in vivo tumor models.[30] These advances in clinical treatment have shown how important sialic acids are in the progression and development of cancer, ultimately indicating that a sialic-acid-blocking drug would be of high therapeutic and clinical treatment value.[31]

Tumor Proliferation and Metastasis:

Expression of sialyl glycans correlated with the aggressive epithelial-mesenchymal transition (EMT), a transition critical in forming metastasis.[32] The EMT process, triggered by growth factor-β (TGF-β), caused the upregulation of numerous sialyltransferases, including ST3GAL1, ST6GAL1, ST8SIA4, ST6GAL2, ST8SIA1, ST8SIA2, and ST3GAL2.[33] When these sialyltransferases upregulate, accumulation of sialoglycans on the cell surface membranes occur, greatly helping tumor cells to survive and metastasize.[34] In hepatocellular carcinoma, the specific α2,6-sialylation activates the Wnt/β-catenin signaling pathway, which results in proliferation, migration, and invasion of cancerous cells.[34]

Through the Akt and ERK pathways, this same α2,6-sialylation, when overexpressed on the human epidermal growth factor receptor 2 (HER2), boosts gastric cancer growth.[35] ST6GAL1 causes hyper-sialylation on the endothelial growth factors receptor (EGFR) through the PI3K/Akt pathway.[36] Likewise, prevention of ST6GAL1's functions can cause induced EGFR to be desialylated and therefore decrease tumor proliferation.[37] The integrin α5β1, modified by α2,6-sialylation, attunes the FAK signaling pathways and cell adhesion.[38] By interrupting the NCAM signaling at the cell-cell communication site, polysialic acid ensures tumor growth and mutations.[39] To slow the rate of metastasis in a xenograft rhabdomyosarcoma tumor in a mouse model, scientists blocked the poly sialyltransferases ST8SIA2 and ST8SIA4 from creating polysialylated NCAM.[40]

As an important observation, sialyl-glycans thickening the surface of tumor cells also helps their colonization and speeds up metastasis.[41] As an example, sialic acid, when liganded with selectins, which are vascular adhesion molecules, magnifies cancer progression by strengthening the adhesion and extravasation processes in metastasis.[42]

DOI: 10.36838/v4i6.14

Tumor Angiogenesis:

Angiogenesis is the creation of more blood vessels using old ones.[49] In the fast-paced environment suitable for metastasis, new blood vessel growth, to provide oxygen and nutrients, is omnipresent.[49] After being synthesized by inflammatory cells, angiogenic growth factors (AGFs) go on to initiate angiogenesis.[49] Compared to other types of AGF, the vascular endothelial growth factor (VEGF) family of VEGF-A, VEGF-B, VEGF-C, VEGF-D, and VEGF-E plays the most significant role.[43] This is because the VEGF undergoes cell-cell communication with polysialic acid.44 In an oxygen-deprived environment, sialic acids assist tumor angiogenesis by upregulating growth factor–receptor interactions.[45]

Gangliosides are a sialyl-glycosphingolipid that can be placed on the endothelial cells' membrane and promote angiogenesis through upregulating responsiveness to AGFs.[46] However, through the transportation of the sialyltransferases ST6GAL1 aided by exosomes, tumor cells can also overexpress the surrounding cells' sialic acid levels.[47] Sialyl Glycans such as N-glycans need to be α2,6-sialylated to engage with VEGF and activate angiogenesis in endothelial cells.[48] The platelet endothelial cell adhesion molecule (PECAM) communicates through homophilic interaction, supported by α2,6-sialylation.[49] While the sialylated PECAM communicates and binds with the VEGFR2 receptor and integrin β3 receptor, the prevention of PECAM-VEGFR2 interaction could be induced by the desialylation of ST6GAL1.[49] This ultimately results in the inhibition of apoptosis- and angiogenesis.[49]

Resistance to Apoptosis and Cancer Cell Therapy:

Cell apoptosis could be prevented by sialic acid through the inhibition of two main processes: the Fas receptor-Fas ligand (FasR-FasL) pathway and the anoikis pathway.[50,53] The FasR-FasL interaction, triggered by T cells and promoted through caspase activation, plays an important role in immune cell homeostasis and the human body's immune maintenance.[50] However, when the FasR is sialylated, it no longer allows the binding of FasL to its binding site, which ultimately inhibits apoptosis.[51] For example, the elevation of α2,6-sialylation on Faso, caused by sialyltransferase ST6GAL1, inhibits the colon carcinoma cells' apoptotic signaling.[51] Furthermore, α2,6-sialylation inhibits the process in which FasR internalizes and cancels the positive feedback loops for FasR-FasL interaction apoptosis.[52]

Another apoptosis pathway is anoikis, which is the detachment of cells off of an extracellular matrix (ECM).[53] Anoikis is a vital pathway that ensures the growth of an adherent-independent cell and the cell's attachment to a mismatched matrix.[53] Resistance to anoikis is often more prominent in cancer cells compared to normal epithelial cells.[53] Attachments of cells to ECM, a process mediated by integrin, is the direct root cause of the upregulation of tumor angiogenic response, which speeds up cancer metastasis.[54] The absence of cell-ECM interactions, due to loss of mediation from integrin, downregulates the phosphorylation (the addition of a phosphate group) of downstream effectors, including FAK, ERK1, PI3-K, and MAP kinases.[55] As a result, this makes the cell susceptible to anoikis.[55] However, the α2,6-hypersialylation in the fibronectin receptor integrin α5β1 could block the binding of galectin-1 and therefore prevent anoikis.[56] Unsurprisingly, α2,6-sialylation showed therapeutic drugs ineffective in many cancers, most likely through the sialylation of the EGFR receptors, which was found to reduce the dimerization of EGFR receptors and increase phosphorylation by them.[57,58] All of this reduces the effectiveness of the tyrosine kinase inhibitors.[59]

Glycosyltransferase and Glycosidase enzymes:

The culmination of multiple glycosylation enzymes, whose role is to assist the distribution of specific glycans on a cell's surface, generates the tumor glyco-code.[60] These additions or removals of a glycan completed by glycosylation enzymes are often the consequence of DNA mutations in the genes responsible for glycan synthesis.[60]

Altered regulation of either sialyltransferase or sialidase enzymes has been linked to hyper-sialylation.[61] Sialyltransferases are a specific type of enzyme that sialylate the termini of glycoconjugates.[62] These enzymes consist of 20 different versions that can all catalyze the attachment of sialic acid to a glycan chain, utilizing different glycosidic linkages including α2-3, α2-6, or α2-8.[62] Sialyltransferases, categorized into four types of ST3Gal, ST6Gal, ST6GalNAc, and ST8SIA, are expressed in a tissue-specific manner.[63]

Sialyltransferases, usually upregulated in cancer, have been correlated with the formation of cancerous antigens.[64] Key sialyltransferases playing a significant role in cancer progression include ST3GAL6, ST6GAL1, ST3GAL4, and ST6GalNAc1/2.[65] ST6GAL1 is responsible for many cancer types and has been implicated as the cause of all the cancer hallmarks.[65] Clinically, ST6GAL1 expression can be used as reliable data to predict responses to EGFR/HER2 inhibitors in cancer cells, specifically ovarian.[65]

In the same sialyltransferases family, ST3GAL4 is also overexpressed in cancer cells, making prognosis less accurate and giving way to the formation of Sialyl Lewis X (sLeX) in gastric carcinoma, which will ultimately end in a metastasis.[66] Besides, ST3GAL6 is also responsible for the synthesis of sLeX and E-selectin ligands.[67] Lately, the homing of bone marrow and the ineffectiveness of multiple myeloma therapy is found to be the direct results of ST3GAL6 activity.[67] Luckily, a study has discovered that this can be inhibited using an E-selectin antagonist called GMI-1271.[68] The sialyltransferases ST6GalNAc1, which catalyzes the sTn antigen responsible in cancer development, is associated with the growth of metastasis.[69] The ST6GALNAC2 has been identified by researchers as a suppressor of metastasis in breast cancer and could be potentially used for treatment with galectin-3 inhibitors.[70]

The loss of terminal sialylation in glycoconjugates is catalyzed by sialidase enzymes, which moderate certain functional molecules' binding sites and play an important role in numerous biological processes.[71] The sialidase enzymes can be found in four specific forms: NEU1, NEU2, NEU3, and NEU4.[72] Sialidase can modify the glycocalyx of cancer cells and, by doing this, might render them more vulnerable to the immune system.[73, 86] Sialidase enzymes, therefore, appear as a promising area for novel therapeutic innovations.[74] For instance, oseltamivir phosphate can target NEU1 in pancreatic, breast,

and ovarian cancers.[75] This method can be used to treat chemoresistant cells which are also drug-sensitive.

Siglecs and Cancer treatment:

The atypical glycosylation found in tumor cells' membranes can lead to novel cell-cell communication with immune cells that leads to the suppression of immune response.[76] The thick layer of sialic acids on the surface of cancer cells has long been recognized for its ability to protect tumors from our body's immune system.[77] With scientists believing that sialic acids play a vital role in the evasion of cancer cells from immune response, recent advances in glyco-tools have opened up new possibilities in the curing of cancer through the immuno-effective pathways.[30] Sialylated glycans on glycoproteins and glycolipids are the key to fitting with the lock of Siglecs, a family of lectins that exist on the membrane of many immune cells.[78] In the tumor microenvironment, this interaction between cancer cell sialic acid and Siglecs can mediate intracellular immune cell response that allows for the evasion of sialic acid from the immune cells.[79] Cancer-promoting glycans, specifically the STn and sialyl T, are a classic example of this.[80] The STn and sialyl T compromise the healthy functions of macrophages and dendritic cells, deaden the natural killer (NK) cells, and diminish the creation of guarding T cells.[81] Therefore, determining the glycoside, which is the distinct aspect of glycan on cancer cells, is pivotal to grasping the mechanism through which glycans aid the evasion of cancer cells from the body's immune system.[82]

Impairing sialic acid's role in communication with Siglecs has become a promising strategy for cancer immunotherapy.[30] Therefore, a reduction of sialic acid poses a huge benefit for our immune system.[82] Emerging research has suggested that loss of tumor sialic acid is positively linked to the loss of Siglecs-sialic acid interaction, which can lead to proficient tumor immunity.[82] If we can block sialic acid from forming interactions with our immune cells, we can allow for more CD8+ T cells, filter out myeloid and regulatory T-cells, and promote cytotoxic T-cells which can kill cancer cells.[30] As discovered previously, the interaction between Siglecs-9, Siglecs-7, Siglecs-10, or Siglecs-15 and their specific ligands has shown how bypassing anti-tumor immunity occurs.[82]

When Siglec-9 matches up with sialylated MUC1, the increase in transformation of monocytes into cancer-promoting macrophages and expression of the checkpoint ligand PD-L1 occurs.[83] Siglec-9 is overexpressed on tumor-eliminating T cells in non-small lung cancer and colorectal and ovarian cancer.[83] Thus, diminishing the sialoglyco-Siglec-9 interaction proves to be a potential enhancement of T-cells activation.[83] Siglec-15 can often be found in tumor-infiltrating myeloid cells and is an important player in cancer immune suppression.[84] Siglec-15 inhibits antigen-specific T-cell responses at the tumor key-and-lock site, which compromises T-cell attacks on cancer cells.[84] Therefore, an anti-Siglec-15 antibody can reallow T-cell effectiveness to carry out tumor immunity.[84] CD24, a type of sialoglycoprotein, is an anti-phagocytic signal that cancer cells show in response to attack from Siglec-10-expressing macrophages.[85] So, blocking CD24-Siglec10 communication can boost the clearance of CD24+ tumors and is a potential immunotherapy plan.[85]

Precise glycocalyx addition in cooperation with antibody-sialidase conjugates has been reported as a potentially highly-valuable avenue for cancer immunotherapy.[86] In this strategy, an antibody induces the sialidase to remove specifically harmful sialic acid, those that are overexpressed, from tumor cell membranes and so open up ways for termination of desialylated cancer cells.[86] Other methods include externally delivering a sialidase enzyme to the tumor and using anti-glycan vaccines to block cancerous glycan-lectin interactions.

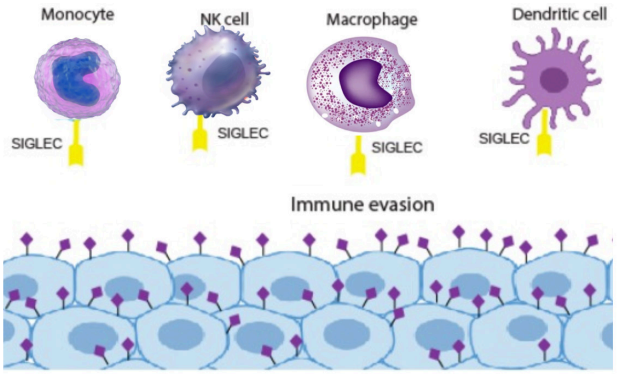


Figure 2: Cell-cell interactions between sialic acid and singles that prevent immunity and promote immune suppression.
Munkley, Jennifer, and Emma Scott. "Targeting Aberrant Sialylation to Treat Cancer." Medicines (Basel, Switzerland) vol. 6,4 102. 14 Oct. 2019, doi:10.3390/medicines6040102

Conclusion

In this review, I covered the key roles sialic acids play in cancer development and progression. A key feature of cancer is hyper-glycosylation, which is the overexpression of sialic acid as terminal sugar on glycans. Sialic acid, whose charge is negative and whose property is hydrophilic, can give the cancer cell many altered traits that help it survive. Over-expressed sialylation is a hallmark of tumors and produces many challenges in curing cancer. As shown above, sialylation is fundamental to the growth, evasion of immunity, progression, and metastasis of cancer cells. Therefore, removing extra sialic acids or inhibiting their functions will likely lead to improvement in the treatment.

Clinical approaches to compromise sialic acids' interactions with Siglecs, thereby canceling the evasion of immunity, proves to be a potentially effective method against cancer development.[87] Playing on the fact that sialic acids hide the glycoproteins' antigenic sites, researchers have been using sialidases to remove sialic acids, so sialic acid can no longer be used as a mask.[87] For instance, when sialidase was present in solid tumor cells, the body's chimeric antigen receptor T-cells (CAR-T cells) successfully removed cancerous cells.[88]

Furthermore, the use of an antibody that titrates singles-15, a protein that binds with sialic acid to suppress T-cells, is expected to be effective in cancer immunotherapy, opening up new advanced treatments for patients whose cells have acquired resistance to the anti-PD-1/PD-L1 therapy.[88] Advances in technologies have led to further study of glycos-

DOI: 10.36838/v4i6.14

ylation, which can ultimately emphasize glycan targeting as the main focus of cancer therapeutics, especially regarding cancer immunotherapy.[88] Shortly, it is probable that the energization of cancer immunotherapy through sialoglycan blocking and traditional chemotherapy/radiation will improve cancer treatment outcomes.[89] Besides being effective drug targets, sialic acids can also be studied to gather more information about cancer cells' sensitivity and resistance to other emerging therapies, which will conclusively improve the fight against cancer.[89]

References

1. The glycocalyx and its significance in human medicine. Tarbell JM, Cancel LM. J Intern Med. 2016 Jul; 280(1):97-113.
2. Biological roles of glycans. Varki A. Glycobiology. 2017 Jan; 27(1):3-49.
3. Varki A., Cummings R.D., Esko J.D., Stanley P., Hart G.W., Aebi M., Darvill A.G., Kinoshita T., Packer N.H., Prestegard J.H., et al. Essentials of Glycobiology. 3rd ed. Cold Spring Harbor Laboratory Press; Cold Spring Harbor, NY, USA: 2017.
4. Zhou, X., Yang, G., & Guan, F. (2020). Biological Functions and Analytical Strategies of Sialic Acids in Tumor. Cells, 9(2), 273. https://doi.org/10.3390/cells9020273
5. Glycosylation in cancer: mechanisms and clinical implications. Pin ho SS, Reis CA. Nat Rev Cancer. 2015 Sep; 15(9):540-55.
6. The role of glycans in the development and progression of prostate cancer. Munkley J, Mills IG, Elliott DJ. Nat Rev Urol. 2016 Jun; 13(6):324-33.
7. The glycosylation landscape of pancreatic cancer. Munkley J. Oncol Lett. 2019 Mar; 17(3):2569-2575.
8. Sialic acid metabolism and sialyltransferases: natural functions and applications. Li Y, Chen X. Appl Microbiol Biotechnol. 2012 May; 94(4):887-905
9. Primary structure and expression analysis of human UDP-N-acetyl-glucosamine-2-epimerase/N-acetylmannosamine kinase, the bifunctional enzyme in neuraminic acid biosynthesis. Lucka L, Krause M, Danker K, Reutter W, Horstkorte R. FEBS Lett. 1999 Jul 9; 454(3):341-4.
10. Identification of the sequence encoding N-acetylneuraminate-9-phosphate phosphatase. Maliekal P, Vertommen D, Delpierre G, Van Schaftingen E. Glycobiology. 2006 Feb;16(2):165-72.
11. Structure and function of vertebrate CMP-sialic acid synthetases. Münster-Kühnel AK, Tiralongo J, Krapp S, Weinhold B, Ritz-Sedlacek V, Jacob U, Gerardy-Schahn R. Glycobiology. 2004 Oct; 14 (10):43R-51R.
12. New Insights into Molecular Organization of Human Neuraminidase-1: Transmembrane Topology and Dimerization Ability.Maurice P, Baud S, Bocharova OV, Bocharov EV, Kuznetsov AS, Kaweckl C, Bocquet O, Romier B, Gorisse L, Ghirardi M, Duca L, Blaise S, Martiny L, Dauchez M, Efremov RG, Debelle L. Sci Rep. 2016 Dec 5; 6():38363.
13. Characterization of human lysosomal neuraminidase defines the molecular basis of the metabolic storage disorder sialidosis. Bonten E, van der Spoel A, Fornerod M, Grosveld G, d'Azzo A. Genes Dev. 1996 Dec 15; 10(24):3156-69.
14. A light-up imaging protocol for neutral pH-enhanced fluorescence detection of lysosomal neuraminidase activity in living cells. Bao L, Ding L, Hui J, Ju H. Chem Commun (Camb). 2016 Oct 25; 52(87) :12897-
15. Cloning and characterization of NEU2, a human gene homologous to rodent soluble sialidases. Monti E, Preti A, Rossi E, Ballabio A, Borsani G. Genomics. 1999 Apr 1; 57(1):137-43.
16. Molecular cloning and characterization of a plasma membrane-associated sialidase specific for gangliosides. Miyagi T, Wada T, Iwamatsu A, Hata K, Yoshikawa Y, Tokuyama S, Sawada M. J Biol Chem. 1999 Feb 19; 274(8):5004-11.
17. Molecular cloning and characterization of NEU4, the fourth member of the human sialidase gene family. Monti E, Bassi MT, Bresciani R, Civini S, Croci GL, Papini N, Riboni M, Zanchetti G, Ballabio A, Preti A, Tettamanti G, Venerando B, Borsani G. Genomics. 2004 Mar; 83(3):445-53.
18. The terminal enzymes of sialic acid metabolism: acylneuraminate pyruvate-lyases. Schauer R, Sommer U, Krüger D, van Unen H, Traving C. Biosci Rep. 1999 Oct; 19(5):373-83.
19. Mechanism of uptake and incorporation of the non-human sialic acid N-glycolylneuraminic acid into human cells. Bardor M, Nguyen DH, Diaz S, Varki A. J Biol Chem. 2005 Feb 11; 280(6):4228-37.
20. Sialic acids sweeten a tumor's life. Büll C, Stoel MA, den Brok MH, Adema GJ. Cancer Res. 2014 Jun 15; 74(12):3199-204.
21. Serum sialylation changes in cancer. Zhang Z, Wuhrer M, Holst S. Glycoconj J. 2018 Apr; 35(2):139-160.
22. Polysialic acid neural cell adhesion molecule (PSA-NCAM) is an adverse prognosis factor in glioblastoma and regulates olig2 expression in glioma cell lines. Amoureux MC, Coulibaly B, Chinot O, Loundou A, Metellus P, Rougon G, Figarella-Branger D. BMC Cancer. 2010 Mar 10; 10():91.
23. Sialic Acid Metabolism: A Key Player in Breast Cancer Metastasis Revealed by Metabolomics. Teoh ST, Ogrodzinski MP, Ross C, Hunter KW, Lunt SY. Front Oncol. 2018; 8():174.
24. Yeo HL, Fan TC, Lin RJ, Yu JC, Liao GS, Chen ES, Ho MY, Lin WD, Chen K, Chen CH, Hung JT, Wu JC, Chang NC, Chang MD, Yu J, Yu AL. Sialylation of vasorin by ST3Gal1 facilitates TGF-β1-mediated tumor angiogenesis and progression. Int J Cancer. 2019 Apr 15;144(8):1996-2007. DOI: 10.1002/ijc.31891.Epub 2019 Jan 3. PMID: 30252131; PMCID: PMC6590135.
25. Expression of ST3GAL4 leads to SLE(x) expression and induces c-Met activation and an invasive phenotype in gastric carcinoma cells. Gomes C, Osório H, Pinto MT, Campos D, Oliveira MJ, Reis CA. PLoS One. 2013; 8(6):e66737.
26. The ST6Gal-I sialyltransferase protects tumor cells against hypoxia by enhancing HIF-1α signaling. Jones RB, Dorsett KA,Hjelmeland AB, Bellis SL. J Biol Chem. 2018 Apr 13; 293(15):5659-5667.
27. E-cadherin and adherens-junctions stability in gastric carcinoma: functional implications of glycosyltransferases involving N-glycan branching biosynthesis, N-acetylglucosaminyltransferases III and V. Pinho SS, Figueiredo J, Cabral J, Carvalho S, Dourado J, Magalhães A, Gärtner F, Mendonca AM, Isaji T, Gu J, Carneiro F, Seruca R, Taniguchi N, Reis CA. Biochim Biophys Acta. 2013 Mar; 1830(3):2690-700.
28. Targeted delivery of a sialic acid-blocking glycomimetic to cancer cells inhibits metastatic spread. Büll C, Boltje TJ, van Dinther EA, Peters T, de Graaf AM, Leusen JH, Kreutz M, Figdor CG, den Brok MH, Adema GJ. ACS Nano. 2015 Jan 27; 9(1):733-45.
29. Targeting aberrant sialylation in cancer cells using a fluorinated sialic acid analog impairs adhesion, migration, and in vivo tumor growth. Büll C, Boltje TJ, Wassink M, de Graaf AM, van Delft FL, den Brok MH, Adema GJ. Mol Cancer Ther. 2013 Oct; 12(10):1935-46.
30. Sialic Acid Blockade Suppresses Tumor Growth by Enhancing T-cell-Mediated Tumor Immunity. Büll C, Boltje TJ, Balneger N, Weischer SM, Wassink M, van Gemst JJ, Bloemendal VR, Boon L, van der Vlag J, Heise T, den Brok MH, Adema GJ. Cancer Res. 2018 Jul 1; 78(13):3574-3588.
31. Munkley, J., & Scott, E. (2019). Targeting Aberrant Sialylation to Treat Cancer. Medicines (Basel, Switzerland), 6(4), 102.https://doi.org/10.3390/medicines6040102

32.Dynamic Sialylation in Transforming Growth Factor-β (TGF-β)-induced Epithelial to Mesenchymal Transition. Du J, Hong S, Dong L, Cheng B, Lin L, Zhao B, Chen YG, Chen X. J Biol Chem. 2015 May 8; 290(19):12000-13.
33.Sialyltransferase ST3GAL1 promotes cell migration, invasion, and TGF-β1-induced EMT and confers paclitaxel resistance in ovarian cancer. Wu X, Zhao J, Ruan Y, Sun L, Xu C, Jiang H. Cell Death Dis. 2018 Oct 30; 9(11):1102.
34.α2,6-Sialylation mediates hepatocellular carcinoma growth in vi tro and in vivo by targeting the Wnt/β-catenin pathway. Zhao Y, Wei A, Zhang H, Chen X, Wang L, Zhang H, Yu X, Yuan Q, Zhang J, Wang S. Oncogenesis. 2017 May 29; 6(5):e343.
35.Increasing HER2 α2,6 sialylation facilitates gastric cancer pro gression and resistance via the Akt and ERK pathways. Liu N, Zhu M, Linhai Y, Song Y, Gui X, Tan G, Li J, Liu Y, Deng Z, Chen X, Wang J, Jia L, He X, Wang X, Lin S. Oncol Rep. 2018 Nov; 40(5):2997-3005.
36.The regulatory ZFAS1/miR-150/ST6GAL1 crosstalk modulates sialylation of EGFR via PI3K/Akt pathway in T-cell acute lym phoblastic leukemia. Liu Q, Ma H, Sun X, Liu B, Xiao Y, Pan S, Zhou H, Dong W, Jia L. J Exp Clin Cancer Res. 2019 May 16; 38(1):199.
37.Enhancement by Nano-Diamino-Tetrac of Antiproliferative Acti-on of Gefitinib on Colorectal Cancer Cells: Mediation by EGFR Sialylation and PI3K Activation. Chang TC, Chin YT, Nana AW, Wang SH, Liao YM, Chen YR, Shih YJ, Changou CA, Yang YS, Wang K, Whang-Peng J, Wang LS, Stain SC, Shih A, Lin HY, Wu CH, Davis PJ. Horm Cancer. 2018 Dec; 9(6):420-432.
38.Caveolin-1 up-regulates integrin α2,6-sialylation to promote inte grin α5β1-dependent hepatocarcinoma cell adhesion. Yu S, Fan J, Liu L, Zhang L, Wang S, Zhang J. FEBS Lett. 2013 Mar 18;58(6) :782-7.
39.A Possible Modulation Mechanism of Intramolecular and Inter-molecular Interactions for NCAM Polysialylation and Cell Migra-tion. Lu B, Liu XH, Liao SM, Lu ZL, Chen D, Troy Ii FA,Huang RB, Zhou GP. Curr Top Med Chem. 2019; 19(25):2271-2282.
40.Intrabodies against the Polysialyltransferases ST8SiaII and ST8-SiaIV inhibit Polysialylation of NCAM in rhabdomyosarcoma tu-mor cells. Somplatzki S, Mühlenhoff M, Kröger A, Gerardy-Schahn R, Böldicke T. BMC Biotechnol. 2017 May 12; 17(1):42.
41.Selectin ligand sialyl-Lewis x antigen drives metastasis of hormone -dependent breast cancers. Julien S, Ivetic A, Grigoriadis A, QiZe D, Burford B, Sproviero D, Picco G, Gillett C, Papp SL, Schaffer L, Tutt A, Taylor-Papadimitriou J, Pinder SE, Burchell JM. Cancer Res. 2011 Dec 15; 71(24):7683-93.
42.Monocyte Induction of E-Selectin-Mediated Endothelial Activa tion Releases VE-Cadherin Junctions to Promote Tumor Cell Ex travasation in the Metastasis Cascade. Häuselmann I, Roblek M, Protsyuk D, Huck V, Knopfova L, Grässle S, Bauer AT, Schneider SW, Borsig L. Cancer Res. 2016 Sep 15; 76(18):5302-12.
43.Mechanisms and regulation of endothelial VEGF receptor signal ling. Simons M, Gordon E, Claesson-Welsh L. Nat Rev Mol Cell Biol. 2016 Oct; 17(10):611-25.
44.Polysialic acid is released by human umbilical vein endothelial cells (HUVEC) in vitro. Strubl S, Schubert U, Kühnle A, Rebl A, Ah-madvand N, Fischer S, Preissner KT, Galuska SP. Cell Biosci. 2018; 8():64.
45.Sialic acid as a target for the development of novel antiangiogenic strategies. Chiodelli P, Urbinati C, Paiardi G, Monti E, Rusnati M. Future Med Chem. 2018 Dec; 10(24):2835-2854.
46.Cell membrane GM1 ganglioside is a functional coreceptor for fibroblast growth factor 2. Rusnati M, Urbinati C, Tanghetti E, Dell'Era P, Lortat-Jacob H, Presta M. Proc Natl Acad Sci U S A. 2002 Apr 2; 99(7):4367-72.
47.Transfer of Functional Cargo in Exomeres. Zhang Q, Higginboth am JN, Jeppesen DK, Yang YP, Li W, McKinley ET, Graves-Deal R, Ping J, Britain CM, Dorsett KA, Hartman CL, Ford DA, Allen RM, Vickers KC, Liu Q, Franklin JL, Bellis SL, Coffey RJ. Cell Rep. 2019 Apr 16; 27(3):940-954.e6.
48.Site-Specific N-Glycosylation of Endothelial Cell Receptor Tyro-sine Kinase VEGFR-2. Chandler KB, Leon DR, Meyer RD, Rahi-mi N, Costello CE. J Proteome Res. 2017 Feb 3; 16(2):677-688.
49.Glycosylation controls cooperative PECAM-VEGFR2-β3 interin functions at the endothelial surface for tumor angiogenesis. Imam aki R, Ogawa K, Kizuka Y, Komi Y, Kojima S, Kotani N, Honke K, Honda T, Taniguchi N, Kitazume S. Oncogene. 2018 Aug; 37(31): 4287-4299.
50.Cell death induced by the Fas/Fas ligand pathway and its role in pathology. Waring P, Müllbacher A. Immunol Cell Biol. 1999 Aug; 77(4):312-7.
51.Sialylation of the Fas death receptor by ST6Gal-I provides protec-tion against Fas-mediated apoptosis in colon carcinoma cells. Swin dall AF, Bellis SL. J Biol Chem. 2011 Jul 1; 286(26):22982-90.
52.Liu, Z., Swindall, A. F., Kesterson, R. A., Schoeb, T. R., Bullard, D. C., & Bellis, S. L. (2011). ST6Gal-I regulates macrophage apop-tosis via α2-6 sialylation of the TNFR1 death receptor. The Jour-nal of biological chemistry, 286(45), 39654–39662. https://doi.org/ 10.1074/jbc.M111.276063Anoikis: cancer and the homeless cell. Liotta LA, Kohn E. Nature. 2004 Aug 26; 430(7003):973-4.
53.Anoikis and EMT: Lethal "Liaisons" during Cancer Progression. Cao Z, Livas T, Kyprianou N. Crit Rev Oncog. 2016; 21(3-4):155-168.
54.Integrin-EGFR interaction regulates anoikis resistance in colon cancer cells. Guha D, Saha T, Bose S, Chakraborty S, Dhar S, Khan P, Adhikary A, Das T, Sa G. Apoptosis. 2019 Dec; 24(11-12):958-971.
55.Targeting constitutively activated β1 integrins inhibits prostate cancer metastasis. Lee YC, Jin JK, Cheng CJ, Huang CF, Song JH, Huang M, Brown WS, Zhang S, Yu-Lee LY, Yeh ET, McIntyre BW, Logothetis CJ, Gallick GE, Lin SH. Mol Cancer Res. 2013 Apr; 11(4):405-17.
56.Emerging role of alpha2,6-sialic acid as a negative regulator of ga-lectin binding and function. Zhuo Y, Bellis SL. J Biol Chem. 2011 Feb 25; 286(8):5935-41.
57.Sialylation of beta1 integrins blocks cell adhesion to galectin-3 and protects cells against galectin-3-induced apoptosis. Zhuo Y, Cham mas R, Bellis SL. J Biol Chem. 2008 Aug 8; 283(32):22177-85.
58.Effect of sialylation on EGFR phosphorylation and resistance to tyrosine kinase inhibition. Yen HY, Liu YC, Chen NY, Tsai CF, Wang YT, Chen YJ, Hsu TL, Yang PC, Wong CH. Proc Natl Acad Sci U S A. 2015 Jun 2; 112(22):6955-60.
59.Glycosylation is an Androgen-Regulated Process Essential for Prostate Cancer Cell Viability. Munkley J, Vodak D, Livermore KE, James K, Wilson BT, Knight B, Mccullagh P, Mcgrath J, Crundwell M, Harries LW, Leung HY, Robson CN, Mills IG, Rajan P, Elliott DJ. EBioMedicine. 2016 Jun; 8():103-116.
60.Sialylation: an Avenue to Target Cancer Cells. Vajaria BN, Patel KR, Begum R, Patel PS. Pathol Oncol Res. 2016 Jul; 22(3):443-7.
61.Sialosignaling: sialyltransferases as engines of self-fueling loops in cancer progression. Dall'Olio F, Malagolini N, Trinchera M, Chiricolo M. Biochim Biophys Acta. 2014 Sep; 1840(9):2752-64.
62.Differential expression of alpha-2,3-sialyltransferases and al pha-1,3/4-fucosyltransferases regulates the levels of sialyl Lewis a and sialyl Lewis x in gastrointestinal carcinoma cells. Carvalho AS, Harduin-Lepers A, Magalhães A, Machado E, Mendes N, Costa LT, Matthiesen R, Almeida R, Costa J, Reis CA. Int J Biochem Cell Biol. 2010 Jan; 42(1):80-9
63.Elevation of β-galactoside α2,6-sialyltransferase 1 in a fructoser

DOI: 10.36838/v4i6.14

esponsive manner promotes pancreatic cancer metastasis. Hsieh CC, Shyr YM, Liao WY, Chen TH, Wang SE, Lu PC, Lin PY, Chen YB, Mao WY, Han HY, Hsiao M, Yang WB, Li WS, Sher YP, Shen CN. Oncotarget. 2017 Jan 31; 8(5):7691-7709.
64.ST6Gal-I overexpression facilitates prostate cancer progression via the PI3K/Akt/GSK-3β/β-catenin signaling pathway. Wei A, Fan B, Zhao Y, Zhang H, Wang L, Yu X, Yuan Q, Yang D, Wang S. Oncotarget. 2016 Oct 4; 7(40):65374-65388.
65.ST6GAL1: A key player in cancer. Garnham R, Scott E, Liver more KE, Munkley J. Oncol Lett. 2019 Aug; 18(2):983-989.
66.Expression of ST3GAL4 leads to SLe(x) expression and induces c-Met activation and an invasive phenotype in gastric carcinoma cells. Gomes C, Osório H, Pinto MT, Campos D, Oliveira MJ, Reis CA. PLoS One. 2013; 8(6):e66737.
67.The sialyltransferase ST3GAL6 influences homing and survival in multiple myeloma. Glavey SV, Manier S, Natoni A, Sacco A, Moschetta M, Reagan MR, Murillo LS, Sahin I, Wu P, Mishima Y, Zhang Y, Zhang W, Zhang Y, Morgan G, Joshi L, Roccaro AM, Ghobrial IM, O'Dwyer ME. Blood. 2014 Sep 11; 124(11):1765-76.
68.E-selectin ligands recognized by HECA452 induce drug resistance in myeloma, which is overcome by the E-selectin antagonist, GMI-1271. Natoni A, Smith TAG, Keane N, McEllistrim C, Connolly C, Jha A, Andrulis M, Ellert E, Raab MS, Glavey SV, Kirkham-McCarthy L, Kumar SK, Locatelli-Hoops SC, Oliva I, Fogler WE, Magnani JL, O'Dwyer ME. Leukemia. 2017 Dec; 31(12): 2642-2651.
69.The androgen receptor controls expression of the cancer-associated sTn antigen and cell adhesion through induction of ST6GalNAc1 in prostate cancer. Munkley J, Oltean S, Vodák D, Wilson BT, Livermore KE, Zhou Y, Star E, Floros VI, Johannessen B, Knight B, McCullagh P, McGrath J, Crundwell M, Skotheim RI, Robson CN, Leung HY, Harries LW, Rajan P, Mills IG, Elliott DJ. Oncotarget. 2015 Oct 27; 6(33):34358-74.
70.An in vivo functional screen identifies ST6GalNAc2 sialyltransferase as a breast cancer metastasis suppressor. Murugaesu N, Iravani M, van Weverwijk A, Ivetic A, Johnson DA, Antonopoulos A, Fearns A, Jamal-Hanjani M, Sims D, Fenwick K, Mitsopoulos C, Gao Q, Orr N, Zvelebil M, Haslam SM, Dell A, Yarwood H, Lord Lord CJ, Ashworth A, Isacke CM. Cancer Discov. 2014 Mar;4(3): 304-17.
71.Mammalian sialidases: physiological and pathological roles in cellular functions. Miyagi T, Yamaguchi K. Glycobiology. 2012 Jul; 22(7):880-96.
72.Sialidase activity in human pathologies. Glanz VY, Myasoedova VA, Grechko AV, Orekhov AN. Eur J Pharmacol. 2019 Jan 5; 842 ():345-350.
73.Sanford BH. An alteration in tumor histocompatibility induced by neuraminidase. Transplantation. 1967 Sep 5;5(5):1273-9. doi: 10.1097/00007890-196709000-00005. PMID: 6056941.
74.Deregulation of sialidases in human normal and tumor tissues. Forcella M, Mozzi A, Stefanini FM, Riva A, Epistolio S, Molinari F, Merlo E, Monti E, Fusi P, Frattini M. Cancer Biomark. 2018 Feb 14; 21(3):591-601.
75.Neuraminidase-1: a novel therapeutic target in multistage tumorigenesis. Haxho F, Neufeld RJ, Szewczuk MR. Oncotarget. 2016 Jun 28; 7(26):40860-40881.
76.The tumour glyco-code as a novel immune checkpoint for immunotherapy. RodrÍguez E, Schetters STT, van Kooyk Y. Nat Rev Immunol. 2018 Mar; 18(3):204-211.
77.Immunogenicity of L 1210 murine leukaemia cells after treatment with neuraminidase. Bagshawe KD, Currie GA. Nature. 1968 Jun 29; 218(5148):1254-5.
78.Siglecs and immune regulation. Pillai S, Netravali IA, Cariappa A, Mattoo H. Annu Rev Immunol. 2012; 30():357-92.
79.Engagement of myelomonocytic Siglecs by tumor-associated ligands modulates the innate immune response to cancer. Läubli H, Pearce OM, Schwarz F, Siddiqui SS, Deng L, Stanczak MA, Deng L, Verhagen A, Secrest P, Lusk C, Schwartz AG, Varki NM, Bui JD, Varki A. Proc Natl Acad Sci U S A. 2014 Sep 30; 111(39):14211-6.
80.Stanczak M.A., Siddiqui S.S., Trefny M.P., Thommen D.S., Boligan K.F., von Gunten S., von Bergwelt-Baildon M. Self-associated molecular patterns mediate cancer immune evasion by engaging Siglecs on T cells. J. Clin. Investig. 2018;128:4912–4923. doi: 10.1172/JCI120612.
81.Interactions between Siglec-7/9 receptors and ligands influence NK cell-dependent tumor immunosurveillance. Jandus C, Boligan KF, Chijioke O, Liu H, Dahlhaus M, Démoulins T, Schneider C, Wehrli M, Hunger RE, Baerlocher GM, Simon HU, Romero P, Münz C, von Gunten S. J Clin Invest. 2014 Apr; 124(4):1810-20.
82.Pearce O.M., Laubli H. Sialic acids in cancer biology and immunity. Glycobiology. 2016;26:111–128. doi: 10.1093/glycob/cwv097
83.The mucin MUC1 modulates the tumor immunological microenvironment through engagement of the lectin Siglec-9. Beatson R, Tajadura-Ortega V, Achkova D, Picco G, Tsourouktsoglou TD, Klausing S, Hillier M, Maher J, Noll T, Crocker PR, Taylor- Papadimitriou J, Burchell JM. Nat Immunol. 2016 Nov; 17(11):1273-1281.
84.Siglec-15 as an immune suppressor and potential target for normalization cancer immunotherapy. Nat. Med. 2019;25:656–666. doi: 10.1038/s41591-019-0374-x.
85.Barkal A.A., Brewer R.E., Markovic M., Kowarsky M., Barkal S.A., Zaro B.W., Weissman I.L. CD24 signalling through macrophage Siglec-10 is a target for cancer immunotherapy. Nature. 2019 ;572:392–396. doi: 10.1038/s41586-019-1456-0.
86.Precision glycocalyx editing as a strategy for cancer immunotherapy. Xiao H, Woods EC, Vukojicic P, Bertozzi CR. Proc Natl Acad Sci U S A. 2016 Sep 13; 113(37):10304-9.
87.Immunotherapy and Hyperprogression: Unwanted Outcomes, Unclear Mechanism. Knorr DA, Ravetch JV. Clin Cancer Res. 2019 Feb 1; 25(3):904-906.
88.King T., Posey A.D. Co-expression of an engineered cell-surface sialidase by CART cells improves anti-cancer activity of NK cells in solid tumors. Cytotherapy. 2019;21:S27. doi: 10.1016/j.jcyt.2019.03.338.
89.Getting a grip on glycans: A current overview of the metabolic oligosaccharide engineering toolbox. Sminia TJ, Zuilhof H, Wennekes T. Carbohydr Res. 2016 Nov 29; 435():121-141

Author

Phuc Nguyen is an immigrant from Vietnam and now a sophomore at La Quinta High School. He aspires to be a neuro-oncologist to find a cure for neuroblastoma. He plans to pursue biomedical sciences in college.

IJHSR
ijhighschoolresearch.org
■ REVIEW ARTICLE

The Effects of Mirror Therapy on Outcomes in the Upper Extremity Post-Stroke

Anush Raghav Polamraju
1345 Easton Avenue, Somerset, New Jersey, 08854, USA; arpolamraju@gmail.com

ABSTRACT: Stroke is the leading cause of disability for adults worldwide, resulting in motor impairments, pain, spasticity, sensory loss, and reduced activities of daily life (ADLs). Mirror therapy (MT) promotes motor function and sensory recovery in patients across stroke acuity. This study analyzes the effects of MT on different levels of impairment in the upper extremities caused by stroke compared to other conventional treatments and other interventions. Sixteen studies (ten RCTs and five systematic reviews, and one metanalysis) were included, consisting of chronic or acute/subacute patients with paresis in their upper extremities. Studies were extracted from PubMed using various keywords: stroke, upper extremity, mirror therapy. MT was significantly more effective than control therapy in motor function and sensory recovery across stroke acuity. No significant differences were found between MT and control treatment in pain, spasticity, and ADLs. MT is associated with improved motor function and sensory recovery in patients with paresis across all stroke acuity.

KEYWORDS: Behavioral and Social Sciences; Neuroscience; Stroke; mirror therapy; upper extremity.

■ Introduction

Background:

Strokes, or cerebrovascular accidents, are a leading cause of death and long-term disability in adults worldwide.[1,2] In the United States, 795,000 individuals experience a stroke each year.[1] Worldwide, strokes affect 15 million individuals.[2] Strokes result from the obstruction or burst of a vessel to the brain, resulting in necrosis, or cell death, in the affected area of the brain.[1-3] Though the effects of a stroke are variable, 85% of stroke survivors experience hemiparesis, or one-sided weakness, in their upper extremities contralateral to the area in which the stroke occurred.[3,4] This can result in movement dysfunction, motor impairments, sensory loss, and reduced quality of life.[2,4] According to the Centers for Disease Control and Prevention, 25% of stroke survivors require assistance or use assistive devices in day-to-day tasks.[1] As a result, it is crucial to research different interventions that affect stroke rehabilitation.

Mirror therapy (MT) is an intervention whose initial purpose was to help amputees recover from phantom limb pain.[4-9] In MT, a mirror is placed in the patient's mid-sagittal plane, reflecting the movements of the non-paretic arm, with the paretic arm being obstructed by the mirror (shown on the figure on the right).[4,6] When the patient watches the mirror, there is a visual illusion that their paretic arm is fully functional, hypothesized to reverse learned nonuse.[5,6] MT has been reported to be effective in patients across stroke acuity with hemiparesis by immediately promoting motor function, motor recovery, and quality of life.[6,9-10]

(Source: EBRSR.org)

The patient observes the mirror reflecting the movements with the non-paretic hand, producing the illusion that the paretic arm is functional:

The purported mechanism behind MT is the activation of the mirror neuron system.[4,6,7,9] During MT, fMRIs have found increased activity in the superior temporal sulcus due to stimulation of the mirror neuron system.[4,7,9] When the patient is directly observing movements, the increased activity in this region causes activation of the primary motor cortex on the lesioned side, resulting in the recovery and activation of motor neurons.[4,7,9] The activation of the primary motor cortex region results in a more significant improvement of motor function and recovery for the patient.[4,7,9]

MT is performed using one of the following three methods:[5]

i) Patients watch the movements made by the non-paretic arm in the mirror while simultaneously imitating movements with the paretic arm.

DOI: 10.36838/v4i6.15

ii) Patients watch the movements made by the non-paretic arm in the mirror while imagining the actions of their affected limb.

iii) Patients watch the movements made by the non-paretic arm in the mirror while being assisted in moving their paretic arm to imitate the actions of the non-paretic arm.

In short, MT might be a simple, cost-effective treatment that patients can receive at home or in a clinical setting to achieve improved arm function.[4,7] New types of MT are being studied, including task-based MT, gesture-recognition MT, and combined MT.[4,11-13]

This review will be utilizing the World Health Organization's International Classification of Function, Disability, and Health, or ICF, model. In this model, the functional abilities of an individual post-disability are assessed mainly in terms of bodily impairments and activities. In the discussion section of this review, the effect of MT on the outcomes of the upper extremities post-stroke will be organized via the ICF model. Ultimately, the objective of this review is to analyze the effects of MT on different levels of impairment in patients' upper extremities post-stroke compared to other stroke interventions and techniques.

Discussion

This section will present the evidence for MT following the ICF model, where outcomes will be divided into impairments, such as motor function loss, pain, spasticity, and somatosensation, as well as activities, such as independence in activities of daily life, in the upper extremities of stroke patients. The section will also discuss variables that may affect MT outcomes, such as dosage and acuity, and introduce variations of MT. Finally, the section will assess the appropriate and inappropriate patients for MT and present several unanswered questions about MT that future studies should address.

Impairments:

Motor function is the speed and excursion of a movement stimulated by the activity of motor neurons.[14] Motor function is usually affected by stroke, resulting in motor impairments in stroke survivors. MT has been reported to be effective in improving motor function in patients compared to control therapy. Five randomized controlled trials (RCTs) examined the effect of MT on motor function and found that there were statistically significant results in the MT group compared to the control group when assessed by the Fugl-Meyer Assessment.[7,10,15-17] This finding was in agreement with one systematic review and one meta-analysis with a total of 73 studies, which found that MT significantly improved motor function immediately after the intervention period compared to control therapy and other interventions.[6,9] Moreover, hemiparesis is a frequent consequence of a stroke associated with functional loss of the upper extremity. There is evidence that MT can improve motor function assessed by the Fugl-Meyer Assessment compared to control therapy in patients with severe hemiparesis.[5,17] In a review examining 15 studies, the authors concluded MT had significant effects on motor function in patients with hemiparesis compared to traditional therapy.[13] These findings suggest that MT possibly enhances motor function in stroke patients with severe motor impairments and hemiparesis.

Pain in the upper extremity post-stroke is also a common symptom for stroke patients. However, MT was found to be ineffective in reducing pain. Compared to control therapy, researchers in one RCT found that patients with severe hemiparesis treated with MT did not have statistically significant outcomes in pain reduction.[5] This finding is consistent with the conclusions of a systematic review including 15 studies that there were no intergroup differences between the MT and control group on pain.[13] The collective results from both the RCT and the review suggest that MT is unlikely to improve pain reduction more significantly than control therapy. However, reduced pain following MT was reported by two systematic reviews for stroke patients with complex regional pain syndrome I (CRPS-I).[6,8] A systematic review examining 62 MT studies with 1,983 patients evaluated pain, measured by the Numerical Rating Scales and the pain section of the Fugl-Meyer Assessment, and found reduced pain for those with CRPS-I post-stroke.[6] Two studies included in the other review found that MT had significant effects on pain for CRPS-I patients.[8] These suggest that MT may be effective in reducing pain in CRPS-I patients post-stroke. Collectively, these results suggest that MT compared to conventional treatment is not effective in pain reduction in stroke patients that are diagnosed with CRPS-I.

Spasticity is another consequence of stroke caused by damage to the upper motor neurons causing increased resistance to passive movement.[18] MT has been reported to be ineffective on spasticity.[4-5,8,15,16] One RCT reported that patients performing MT reported a difference in elbow flexion, wrist flexion, wrist extension, and finger extension between pre-treatment and post-treatment results when measured by the Modified Ashworth Scale.[4] However, there was no statistically significant difference in spasticity scores between the MT and control therapy groups.[4] These findings aligned with the conclusions made in a systematic review containing five RCTs, where it was concluded that groups of patients that performed MT did not demonstrate substantial benefits on spasticity compared to conventional rehabilitation.[8] The use of MT poststroke, therefore, seems to yield no significant improvements in spasticity compared to conventional therapy.

Along with stimulating the primary motor cortex, MT modulates the primary sensory cortex, therefore, somatosensory recovery can occur in the upper extremity.[10] Two RCTs were conducted and examined the effect of MT on promoting somatosensation in the upper extremities. Patients in the first RCT receiving MT had significantly improved temperature and pain sensation versus conventional rehabilitation.[10] In the second RCT, which enrolled severe hemiparetic patients, significant improvement in tactile sensation was observed compared to the control group, but no statistically significant differences were demonstrated between groups when assessed by the Ottingham Sensory Assessment.[17] One RCT in a systematic review evaluating interventions for upper limb sensation recovery post-stroke found that MT where patients experienced improved light touch, thermal, and pressure sen-

sation compared to sham therapy when measured by Quantitative Sensory Testing, largely agreeing with the findings reported by the two RCTs.[20] With these conclusions, it is possible for MT to improve aspects of sensation in the upper extremities post-stroke.

Activities:

Activities of daily life (ADLs) are a patient's ability to perform activities related to personal care including bathing, grooming, eating, and toileting.[19] When a patient suffers a stroke, functional loss affects their upper extremities, reducing their ability to be independent in their lives.[1] As a result, recovering the ability to perform ADLs in stroke patients is a primary concern. Three RCTs evaluated the effectiveness of MT on ADLs and found no significant differences between MT and conventional therapy.[5,7,10] Patients enrolled in the MT group only demonstrated pre-post within-group differences measured by the Modified Barthel Index.[5,7,10] However, there were no significant differences between the MT and control groups for ADLs.[5,7,10] Two recent systematic reviews arrived at different conclusions. A 2019 systematic review reviewing fifteen RCTs had similar results to those of the three RCTs, finding no significant differences between the MT group and the control group.[13] However, a 2018 systematic review concluded that MT had produced a significant effect on ADLs compared to all other interventions.[6] To explain these contrasting findings, the 2018 systematic review included studies in which patients in the experimental group performed task-based MT, which has been concluded to significantly improve ADLs. These results suggest that MT, without modifications, is unlikely to show significant intergroup differences in ADLs compared to control therapy or other interventions.

Dosage:

The dosage of MT reported in various RCTs included in this review ranged from 20 to 90 minutes per session over three to six weeks.[4,5,7,10,11,15-17,21,22] In a systematic review examining 15 studies it was concluded that MT in shorter doses of 30 minutes was more effective in motor function than prolonged doses when compared to conventional therapy.[13] The same systematic review found that the total intervention period also influences the effectiveness of MT in motor recovery of the arm.[13] The review authors concluded that RCTs with an intervention period longer than four weeks were less effective than RCTs that included an intervention period shorter than four weeks.[13] These results suggest that MT should be implemented in shorter doses and intervention periods to improve motor function and recovery.

Acuity:

MT has been reported to be effective on motor function across stroke acuity. In four RCTs, patients who were either chronic or acute/subacute and treated with MT had statistically significant effects on motor function compared to the control group therapy, therefore, concluding that MT can enhance motor function across stroke acuity.[5,7,10,15] Three systematic reviews evaluating studies including patients across stroke acuity also concluded that MT significantly improved motor function measured by the Fugl-Meyer Assessment.[6,9,13] Collectively, the findings suggest that MT can enhance motor function across stroke acuity.

Conversely, reports of the effect of MT on pain differ based on stroke acuity. Patients with chronic stroke who received MT did not experience significant effects on pain compared to control therapy.[5] A systematic review examining fifteen RCTs (fourteen RCTs with chronic subjects) concluded no significant differences in pain between the MT and conventional therapy groups.[13] The remaining RCT enrolled acute patients and reported improved pain in MT compared to conventional therapy.[13] These findings suggest that MT may effectively treat pain in acute but not in chronic or subacute stroke.

Stroke acuity also plays a role in the effectiveness of MT in sensory recovery. One RCT that enrolled chronic patients found improved temperature, tactile, and pain sensation in the MT group compared to the control group.[10] One RCT in a 2010 systematic review that enrolled acute patients found the same improvements in pressure, tactile, and thermal sensation as the RCTs which enrolled chronic patients.[20] These findings suggest that MT can improve aspects of sensory recovery in both acute and chronic stroke.

Patient Appropriateness/Inappropriateness:

To treat patients with MT, it is essential to understand whether MT is a suitable treatment for each patient. Cost, travel, and the necessity for sophisticated, state-of-the-art equipment are not significant issues, as MT utilizes cost-effective, accessible equipment that can be implemented in any setting.[4,7,11] As a result, patients with cost barriers will still be able to perform MT.

To perform MT, patients must follow directions from clinicians and complete the actions/tasks they are prescribed. As a result, patients with receptive aphasia, severe attention deficits, and cognitive impairments may be inappropriate to receive MT as these conditions may interfere with their ability to understand and follow instructions.[4,6-8,11] The Mini-Mental State Exam was utilized in studies to determine whether a patient was likely to participate in MT successfully.[5,6,10,11,15-17,22]

In addition, patients with severe visuospatial neglect and visual impairments may not be successful with MT treatments.[4,6] Since MT requires patients to look at a mirror to perform actions and tasks, patients who cannot clearly see their surroundings will be unable to see the mirror. As a result, these patients are unsuitable to receive MT.

Variations of MT:

New versions of MT are being developed to increase the effectiveness of MT. One such development is task-based MT.[4,21,22] Task-based MT requires the patient to perform goal-oriented practices and functional tasks using the mirror.[4,21,22] These practices require the patient to manipulate everyday objects such as chopsticks or a ball to perform various actions such as wrist extension and elbow flexion.[4,21,22] Two RCTs reported improved motor function and motor recovery of the hand compared to conventional treatment in patients with severe hemiparesis and upper limb impairment.[4,21] One RCT reported task-based MT had statistically significant improvement in ADLs than sham therapy when assessed by the Modified Barthel Index.[22] These results suggest that task-based MT effectively improves motor function and ADLs.

DOI: 10.36838/v4i6.15

Likewise, another development of MT is gesture-recognition MT. Patients performing gesture-recognition MT use a Leap Motion controller, a device that detects motion, and observe at a mirror reflecting the screen of a monitor.[11] Patients in gesture-recognition MT play various game programs that require them to perform multiple actions such as lifting their hands and picking up objects.[11] Gesture-recognition MT was reported to significantly affect motor function and quality of life in an RCT that included chronic patients compared to MT and control therapy.[11]

Collectively, these variations of MT have positive effects on different areas in the upper extremity post-stroke. Task-based MT was found to affect motor function, recovery, and ADLs positively compared to other interventions. Likewise, gesture-based MT also had positive effects on motor function and quality of life.

Unanswered Questions:

Due to its maneuverability, MT can be utilized in any setting, such as at a patient's home or in a clinician's office.[4,7] Multiple systematic reviews have identified this fact and used studies that included at-home MT.[6,13] However, no differences between at-home MT compared to MT in a clinical setting were studied. A possible hypothesis would be that clinical MT under a therapist's supervision will be more effective in improving impairment levels than at-home MT. Future studies should assess the difference between both methods of MT, by separating two groups based on clinical MT and at-home MT. Measuring the patients at baseline, post-intervention, and follow-up, studies should observe differences in motor function, spasticity, ADLs, and somatosensation to determine any intergroup differences and sustained effects of the treatment. Finding any significant differences would determine the setting that will allow stroke patients to receive the most benefit from MT.

Determining an optimal dosage for MT that provides the maximum benefit is a limitation in studies, as multiple studies use different dosage amounts.[5,13] In a systematic review, it was found that the longer the dosage of MT, the effectiveness of MT is reduced.[13] Regarding the conclusions made in the systematic review, it could be hypothesized that shorter doses of MT increase the effectiveness of the treatment. Future studies should research the dosage of MT that should be administered to stroke patients to determine the dosage that will result in the maximum efficacy of MT. Studies should separate patients into groups, each performing at different dosage amounts, and then measure each group at baseline and post-intervention on a variety of outcomes such as motor function, spasticity, ADLs, and sensation recovery. Obtaining any statistically significant differences between all three groups can aid clinicians in determining an optimal dosage amount for MT, which will produce the maximum rehabilitation effect for a patient post-stroke.

Usually, stroke patients seek rehabilitation services to enable them to complete basic day-to-day activities.[1,19] Earlier in this review, it was noted that MT has a possibility of improving different impairment levels such as motor function and somatosensation in the upper extremities. However, MT did not yield significant improvements in an individual's ability to partake in ADLs.[5-10,15-17,20] As a result, future studies should answer whether MT is an intervention that should be recommended by clinicians to use for the general stroke patient population. Future studies should also research variations of MT such as task-based MT as it has been shown to improve both impairment and activity, particularly motor function, somatosensation, and ADLs, in the upper extremities of stroke patients.[4,21,22] Future studies should further study these variations by comparing groups with task-based MT, MT, and conventional treatments in outcomes such as motor function, spasticity, somatosensation, and ADLs, observing for any statistically significant intergroup differences post-intervention. This will allow clinicians to recommend MT to stroke patients and expect positive results in the upper extremities of their patients because of the treatment.

Conclusion

MT has been shown to improve motor function and aspects of sensorimotor recovery compared to conventional therapy across all stroke acuity.[5-7,9,10,13,15-17] No significant differences were reported between conventional treatment and MT in pain, spasticity, or ADLs post-intervention.[4-8,10,13,15-16] However, acute patients and patients with CRPS-I reported less pain than conventional therapy.[6,8,13] It is also important to note that patients with severe hemiparesis had improved motor function and sensory recovery from MT.[5,17] These results suggest that patients with motor impairments, sensory loss, and hemiparesis across all acuity can benefit from MT and, therefore, have it prescribed by a physical therapist.

This review also analyzed several variations of MT, such as task-based MT and gesture-recognition MT. Increased ability to perform ADLs assessed by the Modified Barthel Index, were found in patients performing task-based MT compared to sham therapy.[22] In addition, task-based MT was found to improve motor function compared to conventional treatment more effectively.[21,22] Gesture-recognition MT also had improved motor function compared to traditional MT and conventional therapy.[11] Furthermore, improved quality of life was reported in patients performing gesture-recognition MT compared to conventional treatment.[11] These findings suggest that new developments of MT are also viable options for stroke patients.

Future studies should address the limitations found in this review. Studies should compare at-home MT to clinic MT to determine which form of MT is more effective. If a future study found that MT performed in a home setting is more effective than at a clinic, the impact would be significant. Patients will not need to pay for a therapist and can experience improved results in various outcomes at home compared to being in a clinic. As a result, future studies in this area are crucial due to their immense impact on patients. In terms of dosage, studies should research an optimal dosage of MT that produces the maximum efficacy for a patient. This study showed that shorter doses increase the effectiveness of MT, however, researchers should further study if this conclusion is valid. This would help future clinicians determine the proper dosage of MT to help more patients affected by stroke.

Lastly, studies should definitively conclude the worth of MT as a rehabilitation method for stroke patients by further studying developments of MT that provide positive results in both impairment and activity measures. This would ultimately help clinicians recommend MT to a broader scope of stroke patients and expect positive results in improving impairment levels and activity independence, allowing patients to lead functional lives post-stroke. Moreover, more rigorous, multi-site RCTs with large sample sizes should be conducted to definitively conclude the effectiveness of MT for stroke patients.

Acknowledgments

This research project was supported by my friends, family, teachers, and mentors. Firstly, I would like to thank Dr. Jane E. Sullivan of Northwestern University for providing me with the knowledge to conduct this research, as well as providing me with comprehensive suggestions to improve my paper. Next, I would like to thank my teachers for teaching me the beauty of scientific discovery and research. Lastly, I especially thank my parents, whose support has shaped me into the person I am today and enabled me to pursue my passion for scientific research.

References

1. Stroke facts. https://www.cdc.gov/stroke/facts.htm.
2. Buford, J. A.; Kloos, A. D.; Basso, D. M.; Kegelmeyer, D. A.; Nichols-Larsen, D. S.; Heathcock, J. C. *In Neurologic rehabilitation: Neuroscience and neuroplasticity in Physical therapy practice*; Mc Graw-Hill Education: New York, 2016; pp 151–151.
3. For patients and families. https://strokengine.ca/en/resources/for-patients-and-families/#living-with-stroke. (Accessed Apr 1, 2022).
4. Madhoun, H. Y.; Tan, B.; Feng, Y.; Zhou, Y.; Zhou, C.; Yu, L., Task-based mirror therapy enhances the upper limb motor function in subacute stroke patients: a randomized control trial. *Eur J Phys Rehabil Med* **2020**, 56 (3), 265-271.
5. Thieme, H.; Bayn, M.; Wurg, M.; Zange, C.; Pohl, M.; Behrens, J. Mirror Therapy for Patients with Severe Arm Paresis after Stroke – a Randomized Controlled Trial. *Clinical Rehabilitation* **2012**, 27 (4), 314–324.
6. Thieme, H.; Morkisch, N.; Mehrholz, J.; Pohl, M.; Behrens, J.; Borgetto, B.; Dohle, C. Mirror Therapy for Improving Motor Function after Stroke. *Cochrane Database of Systematic Reviews* **2018**, 2018 (7).
7. Invernizzi, M.; Negrini, S.; Carda, S.; Lanzotti, L.; Cisari, C.; Baricich, A., The value of adding mirror therapy for upper limb motor recovery of subacute stroke patients: a randomized controlled trial. *Eur J Phys Rehabil Med* **2013**, 49 (3), 311-7.
8. Ezendam, D.; Bongers, R. M.; Jannink, M. J., Systematic review of the effectiveness of mirror therapy in upper extremity function. *Disabil Rehabil* **2009**, 31 (26), 2135-49.
9. Zeng, W.; Guo, Y.; Wu, G.; Liu, X.; Fang, Q. Mirror Therapy for Motor Function of the Upper Extremity in Patients with Stroke: A Meta-Analysis. *Journal of Rehabilitation Medicine* **2018**, 50 (1), 8–15.
10. Wu, C.-Y.; Huang, P.-C.; Chen, Y.-T.; Lin, K.-C.; Yang, H.-W. Effects of Mirror Therapy on Motor and Sensory Recovery in Chronic Stroke: A Randomized Controlled Trial. *Archives of Physical Medicine and Rehabilitation* **2013**, 94 (6), 1023–1030.
11. Choi, H. S.; Shin, W. S.; Bang, D. H., Mirror Therapy Using Gesture Recognition for Upper Limb Function, Neck Discomfort, and Quality of Life After Chronic Stroke: A Single-Blind Randomized Controlled Trial. *Med Sci Monit* **2019**, 25, 3271-3278.
12. Luo, Z.;Zhou, Y.;He, H.;Lin, S.;Zhu, R.;Liu, Z.;Liu, J.; Liu, X.; Chen, S.; Zou, J.;Zeng, Q., Synergistic Effect of Combined Mirror Therapy on Upper Extremity in Patients With Stroke: A Systematic Review and Meta-Analysis. *Front Neurol* **2020**, 11, 155.
13. Cantero-Tellez, R.; Naughton, N.; Algar, L.; Valdes, K., Outcome measurement of hand function following mirror therapy for stroke rehabilitation: A systematic review. *J Hand Ther* **2019**, 32 (2), 277-291 e1.
14. Apa Dictionary of Psychology. https://dictionary.apa.org/motor-function.
15. Samuelkamaleshkumar, S.; Reethajanetsureka, S.; Pauljebaraj, P.; Benshamir, B.; Padankatti, S. M.; David, J. A. Mirror Therapy Enhances Motor Performance in the Paretic Upper Limb after Stroke: A Pilot Randomized Controlled Trial. *Archives of Physical Medicine and Rehabilitation* **2014**, 95 (11), 2000–2005.
16. Yavuzer, G.; Selles, R.; Sezer, N.; Sütbeyaz, S.; Bussmann, J. B.; Köseoğlu, F.; Atay, M. B.; Stam, H. J. Mirror Therapy Improves Hand Function in Subacute Stroke: A Randomized Controlled Trial. *Archives of Physical Medicine and Rehabilitation* **2008**, 89 (3), 393–398.
17. Colomer, C.; NOé, E.; Llorens, R. Mirror therapy in chronic stroke survivors with severely impaired upper limb function: A randomized controlled trial. https://www.ncbi.nlm.nih.gov/pubmed/26933644.
18. Apa Dictionary of Psychology. https://dictionary.apa.org/spasticity.
19. Apa Dictionary of Psychology. https://dictionary.apa.org/activities-of-daily-living.
20. Doyle, S.; Bennett, S.; Fasoli, S. E.; McKenna, K. T., Interventions for sensory impairment in the upper limb after stroke. *Cochrane Database Syst Rev* **2010**, (6), CD006331.
21. Arya, K. N.; Pandian, S.; Kumar, D.; Puri, V., Task-Based Mirror Therapy Augmenting Motor Recovery in Poststroke Hemiparesis: A Randomized Controlled Trial. *J Stroke Cerebrovasc Dis* **2015**, 24 (8), 1738-48.
22. Lim, K.-B.; Lee, H.-J.; Yoo, J.; Yun, H.-J.; Hwang, H.-J. Efficacy of Mirror Therapy Containing Functional Tasks in Poststroke Patients. *Annals of Rehabilitation Medicine* **2016**, 40 (4), 629.
pyruvate-lyases. Schauer R, Sommer U, Krüger D, van Unen H, Traving C. Biosci Rep. 1999 Oct; 19(5):373-83.

Author

Anush Polamraju is currently a senior at Rutgers Preparatory School in New Jersey. His research is focused on the rehabilitation of stroke patients and neuroplasticity. He hopes to pursue a neuroscience and biology double-major at university.

DOI: 10.36838/v4i5.15

IJHSR
ijhighschoolresearch.org
REVIEW ARTICLE

A Synopsis of Diagnostic Testing for COVID-19

Siddhartha Pulukuri
Sharon High School, 181 Pond Street, Sharon, Massachusetts, 02067, USA; spulukuri2022@gmail.com

ABSTRACT: Diagnosis is the precursor to the treatment of a disease, and as such, it is vital for diagnosis to be both accurate and fast. Handling large scale outbreaks and pandemics can be tied to the effectiveness of diagnosis. In the current pandemic of COVID-19, both serological tests and RT-PCR are being used to diagnose COVID-19. SARS-CoV-2 initially broke through in Wuhan, China in December 2019, and since then it has rapidly developed into a pandemic with roughly fourteen million cases and roughly one million fatalities as of July 2021. In this scenario, serological testing is fast, and RT-PCR is accurate; however, what we need to effectively control the pandemic is to have a test that embodies both speed, accuracy, and accessibility. In this review, we will delve into the current SARS-CoV-2 diagnostic tests: RT-PCR and serological assays. Due to the novelty of research in the field, we will also discuss the significance of variants and how they affect the tests based on the key mechanisms behind these detection assays.

KEYWORDS: Translational Medical Sciences; COVID-19; RT-PCR; Serological Testing; Disease Prevention and Diagnosis.

Introduction

In December 2019, an idiopathic respiratory pathogen was identified as the culprit behind a cluster of pneumonia-like cases in Wuhan, a city in the Hubei province of China. Soon, the disease would spread rapidly, escalating from an outbreak to an epidemic, to a pandemic.[1] The disease was designated as coronavirus disease 2019 (COVID-19), previously known as 2019 novel coronavirus disease (2019-nCoV), and the virus known as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) by the World Health Organization (WHO) in February 2020.

SARS-CoV-2 is quite similar to SARS-CoV due to their common ancestry; both have virion sizes ranging from 70 to 90 nm as well as spike, membrane, and envelope surface viral proteins that are embedded in a host membrane-derived lipid bilayer encapsulating the helical nucleocapsid comprising of viral RNA (Figure 1A).[2] The SARS-CoV-2 binds to the angiotensin converting enzyme 2(ACE2) receptors on the host cells inside the respiratory tract and is transmitted aerially (Figure 1B & 1C).[3] An interesting point is that the virus doesn't affect people in the same manner, some can become silent carriers who exhibit no symptoms but can spread the virus whereas others may succumb to acute respiratory distress syndrome (ARDS) and these chances can be increased or decreased based on pre-existing conditions such as asthma, immunodeficiency, obesity, and diabetes.[4]

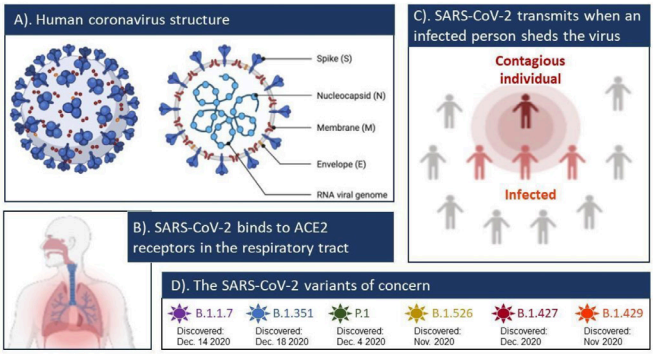


Figure 1: Key features, mode of transmission and variants of SARS-CoV- 2

Over the course of the COVID-19 pandemic, the SARS-CoV-2 virus has mutated, resulting in genetic variation in the circulating viral strains. In Figure 1D, key mutations of the spike protein are shown; mutations can occur on all spike protein subunits, and these variations can enhance the infectivity or virulence.[5] The Delta variant is one of the most recent variants that has been identified and spread significantly. Some potential changes include an increased contagious window, increased viral shedding, increased environmental stability, and increased binding to host ACE2 receptors.

Increased Contagious Window :

The SARS-CoV-2 virus is only infectious when the virus is proliferating in the upper respiratory tract. Genetic variants have the potential to increase the length of these incubation periods, thus increasing the contagious window, allowing for the host to have more chances to pass on the virus.[6]

DOI: 10.36838/v4i6.16

Increased Viral Shedding:

The primary mechanism of transmission for SARS-CoV-2 is droplet transmission via breathing, speaking, coughing, and sneezing. Certain variants can increase the amount of virus that can be shed in these droplets, increasing the viral load as well as the likelihood of passing on the infection.[6]

Increased Environmental Stability:

Often times, hosts can spread infectious droplets onto surfaces in an environment, and it's possible for infection to occur across contact on open surfaces. Variants can prolong the survival of the virus in aerosol droplets, allowing for potential transmission.[6]

Increased binding to host receptor:

The primary method of SARS-CoV-2 infection is by binding to the ACE2 receptor in respiratory tract cells; variants of concern can potentially bind to ACE2 more effectively increasing infectivity.[7]

Latest research regarding SARS-CoV-2 genome, structure, and variants have provided a strong foundation for diagnosis, treatment, and vaccine studies. Diagnosis is a crucial part of treating any patient infected with a disease, and the importance of diagnosis is amplified in situations such as a pandemic. Without a proper diagnosis, treatments can't be administered, and carriers of the disease can slip by undetected, leading to further infection.[8] During this COVID-19 pandemic, diagnosis is the key to managing the pandemic. Currently, labs and testing facilities can't cope with the demand for testing, and faster methods of testing have their own drawbacks in terms of accuracy.[9] In this review, I will discuss the two main forms of diagnosis and their pros and cons.

Discussion

Serological Diagnostic Testing:

Serological tests, or antibody tests, are a form of diagnostic testing being utilized for the diagnosis of COVID-19. With the majority of Serological tests in the US being categorized as Rapid Diagnostic Tests (RDTs), speed distinguishes this form of testing from others. This type of test is known as a Laminar Flow Immunoassay (LFI), meaning that it uses the principle of capillarity, the ability of liquids to flow in narrow spaces without needing energy, to test for the presence of antibodies that indicate a current or previous infection of COVID-19.[10]

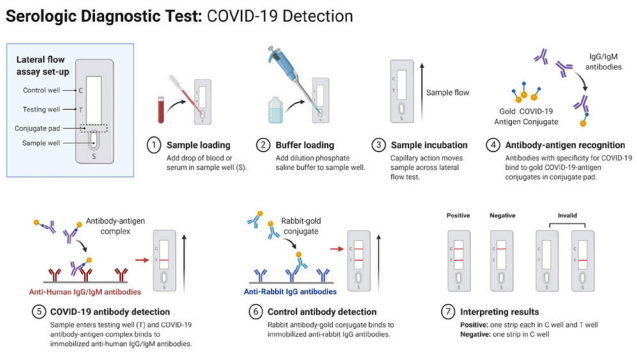


Figure 2: Serologic diagnostic test: COVID-19 detection

Before we can delve into the mechanics of the test, we must first notice its components (Figure 2). The serological diagnostic test set-up consists of four parts: the control well, the testing well, the conjugate pad, and the sample well. Primarily, a drop from the sample of blood or serum is placed on the sample well along with dilute PBS (phosphate-buffered saline).[11] The sample then flows from the sample well to the control well because of capillary action. Afterward, antibodies with specificity for COVID-19 bind to gold COVID-19 antigen conjugates in the conjugate pad. Once the sample manages to flow to the testing well, the COVID-19 antigen-antigen complex binds to immobilized anti-human IgG and IgM antibodies.[11] Promptly after this process, the sample moves to the control well where Rabbit antibody-gold conjugate binds to immobilized anti-rabbit IgG antibodies. Once this phase is complete, you can interpret the results based on the strips. If both control and testing wells have strips, the result is COVID positive, if the control well strip is present but nothing in the testing well, the result is negative, and lastly if the control well doesn't have a strip, the test is invalid.[12]

Although the test focuses on speed, accuracy can sometimes prove to be compromised; for example, when 58 people test positive for COVID-19, of these, 12 people (21%) will not have COVID-19 (false-positive result) and when 942 people test negative for COVID-19, of these, 4 people (0.4%) will actually have COVID-19 (false-negative result). Some tests can give a false positives if the patient the sample was taken from was exposed to a virus similar to SARS-Cov2.[13] Additionally, the test can give a false negative if the patient has an active infection but no antibodies are produced.[14] Lastly, antibody testing promotes a false sense of immunity as people assume that a presence of antibodies means protection from a potential re-infection; however, there are no current studies or evidence that shows that people who have contracted COVID-19 will not be infected again.[15]

RT-PCR Testing:

Reverse Transcriptase Polymerase Chain Reaction (RT-PCR) is a form of diagnostic testing for COVID-19 that focuses on producing accurate results in a timely manner. RT-PCR tests generally take around 3 hours, with unconfirmed tests by some companies clocking in as little as 15 minutes.[16]

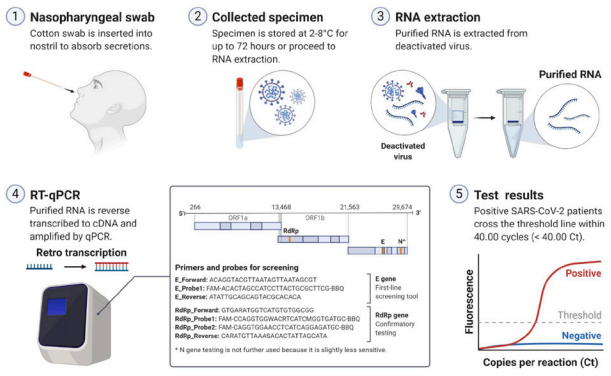


Figure 3: COVID-19 diagnostic test through RT-PCR

DOI: 10.36838/v4i6.16

RT-PCR testing consists of five key steps. As seen in Figure 3, to use RT-PCR, a patient sample is taken via a nasopharyngeal swab and from there it can be stored at 2 to 8 degrees Celsius for up to 72 hours, or it can be taken directly for RNA extraction.[17] During RNA extraction, the sample is placed in a special solution so as to extract only pure RNA from both the virus and from the patient's own cells. To figure out whether or not the virus is present, we must amplify the viral RNA to a point where it is detectable by qPCR. During RT-PCR, the pure RNA is first turned into cDNA (complementary DNA) and the cDNA is in turn amplified by PCR.[17] The test utilizes two primers and a probe to detect two regions in the SARS-CoV-2 RNA-dependent RNA polymerase (RdRp) gene and one primer and a probe set to detect the virus envelope (E) gene in a clinical sample. The machine utilizes enzymes and special chemicals along with heating and cooling cycles to amplify targeted genetic material in the test tube. After several cycles, millions of copies of the SARS-CoV-2 virus's DNA are present in the test tube. While new copies of viral DNA are created, marker labels attach to the DNA strands and release fluorescent dye which can be detected. Once the amount of dye detected reaches a certain value called the "threshold", that is when our testing is complete. If the threshold is crossed in under 40 cycles, the patient has COVID-19, and this number of cycles needed to cross the threshold is called the cycle threshold (Ct).[17]

A positive test indicates an infection, but a negative test indicates that you don't have an infection at the time of the test; however, you could develop an infection and symptoms later. Although the test is very accurate, if the SARS-Cov2 virus were to undergo any mutation throughout its infection in the patient, the RT-PCR may not function as the primers will not be able to bind to the viral RNA due to mismatches. Additionally, the quality of the sample taken directly affects the testing. If inadequate sampling techniques are used, the test will not provide accurate results.[18]

Kinetics of SARS-CoV-2 Positivity of Infected and Recovered Patients:

The kinetics of the PCR and antibody test positivity, represent the variation over time in diagnostic tests for detection of SARS-CoV-2 infection (Figure 4). PCR tests are positive relatively faster in the disease course, whereas serological tests are positive later in the disease course.[19] Due to the sensitivity of PCR, it can detect minuscule amounts of viral material starting from 2-3 days post-infection. The test can continue to detect fragments of the SARS-CoV-2 virus after COVID-19 convalescence that are not infectious. You may continue to test positive if you've had COVID-19 in the distant past, even though you can't spread the SARS-CoV-2 virus to others. Prolonged infection in immunocompromised individuals can occur where they transmit infectious droplets for months.[19] The use of the serologic test for SARS-CoV-2 requires a general understanding of the dynamics of the immune response to infection and specific knowledge of test characteristics. According to the CDC, IgG and IgM are detectable 1-2 weeks post-infection and begin to fall roughly 4-5 weeks post-infection. Although IgG antibodies can stay relatively longer, the true nature of the antibodies detectability hasn't been fully explored yet. Because antibodies require time to become detectable, serologic tests aren't useful early in the course of illness for diagnosing COVID-19. Additionally, most, but not all patients with SARS-CoV-2 infection develop an antibody response, and so negative serologic result does not exclude past infection. Serologic testing in SARS-CoV-2 infection may be useful for identifying symptomatic people suspected of long COVID/post-COVID who may be beyond the period where viral RNA is detectable by RT-PCR.[21]

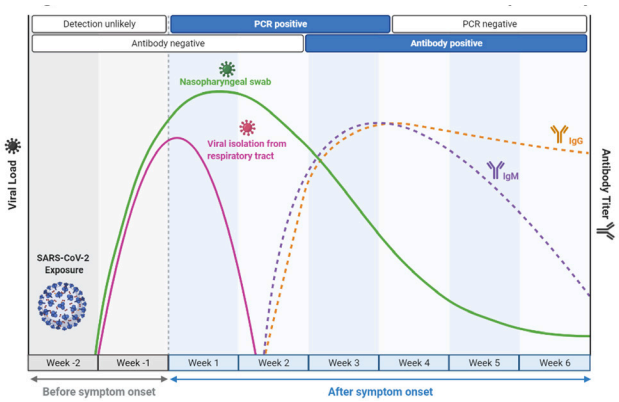


Figure 4: Time course of COVID-19 infection and test positivity

Impact of SARS-CoV-2 Variants:

The impact of variants on the performance of diagnostic tests hasn't been clearly discovered. The majority of commercial assays are designed to detect antibodies to wild-type spike proteins, therefore there's reason to believe that certain variants may generate antibodies undetectable by these assays.[22] However, to date, there are no studies that support variants change analytic or clinical sensitivity in antibody assays. On the other hand, the variants can negatively affect RT-PCR.[23] Researchers hypothesized that mutations in the loci recognized by DNA primers may reduce the amplification of viral sequences which would decrease the efficiency of tests in positive individuals. New studies indicate that diagnostic tests should include multiple target areas to increase the probability of accurate detection. At Washington University, the molecular diagnostic lab of Barnes-Jewish Hospital uses the Roche Cobas SARS-CoV-2 for patient samples. The test utilizes both the ORF1ab as well as the E gene, which encodes the envelope protein. Although most samples have similar cycle threshold values, there were a few outliers: the E gene wasn't amplified to the same degree as ORF1ab. Despite this unexpected result, the tests still performed exceptionally in correctly identifying positive samples based on the ORF1ab signal.[24] To understand the odd results behind the E gene, they sequenced a handful of viral sequences. They found three samples had a common mutation in the E gene, one not present in any of the common variants now circulating in the population. The researchers stated that the mutation affects the binding of the PCR primer which interferes with amplification. As of now, the scientific community requires more studies to test the sensitivity and efficiency of SARS-CoV-2 primer sets.[24]

Conclusion

The US Food and Drug Administration (FDA) has acknowledged that SARS-CoV-2 mutations can interfere with diagnostic tests, and as a result, they're continuing to monitor variants and evaluate potential effects on diagnostic tests. The agency also advised developers to include multiple genetic targets, and they also advised developers to watch for mutations that alter test performance.[25] "From the very beginning, we've been keeping a very close eye on all the different variants, "says Palani Kumaresan, the head of research and development for Roche Diagnostic Solutions. The E gene signal is in fact pan-sarbecovirus, the virus subgenus that includes both SARS-CoV and SARS-CoV-2. SARS-CoV-2, in particular, emits a specific signal from ORF1ab.[26] The US Centers for Disease Control and Prevention (CDC) regularly monitors the primers and probes for its COVID-19 diagnostic panel and multiplex test for flu and COVID-19. CDC also actively tracks and characterizes coronavirus variants through genomic surveillance efforts.

In retrospect, serological tests are great for rapid diagnosis, while RT-PCR is an extremely accurate method of diagnosis.[27] Despite the individual strengths of RT-PCR and serological testing, a test that is both fast and accurate, as well as affordable, is required.[28] Certain companies such as Sorrento Therapeutics, are offering solutions to this problem. This test gives a read-out within 30 minutes of testing a saliva sample, and unlike other tests available, all materials come in a single tube and require no specialized equipment, making it ideal for testing in any situation, potentially even at home.[29] Tests like these, which require no intermediary step between reference labs and facilities collecting samples, can easily overcome the current problem of overwhelmed labs unable to keep up with the demands of testing. Tests such as these have the potential to facilitate the control of the outbreak, as the more accurate diagnoses we have, the easier it becomes to treat and quarantine infected individuals. Hence, it is of paramount importance that we push such tests that have both accuracy and speed into the market.[30]

Acknowledgments

I would like to take the time to acknowledge primarily my family who has supported me in all my scientific endeavors, and also Dr. Ajit Bharti for allowing me to kindle my passion for science in his lab and teaching me the skills to approach problems scientifically.

References

1. Rauf, A., Abu-Izneid, T., Olatunde, A., Ahmed Khalil, A., Alhumaydhi, F.A., Tufail, T., Shariati, M.A., Rebezov, M., Almarhoon, Z. M., Mabkhot, Y.N. and Alsayari, A., 2020. COVID-19 pandemic: epidemiology, etiology, conventional and non-conventional therapies. International Journal of Environmental Research and Public Health, 17(21), p.8155.
2. Wang, M.Y., Zhao, R., Gao, L.J., Gao, X.F., Wang, D.P. and Cao, J.M., 2020. SARS-CoV-2: structure, biology, and structure-based therapeutics development. Frontiers in cellular and infection microbiology, 10.
3. Shi, R., Shan, C., Duan, X., Chen, Z., Liu, P., Song, J., Song, T., Bi, X., Han, C., Wu, L. and Gao, G., 2020. A human neutralizing anti body targets the receptor-binding site of SARS-CoV-2. Nature, 584(7819), pp.120-124.
4. Tzotzos, S.J., Fischer, B., Fischer, H. and Zeitlinger, M., 2020. Incidence of ARDS and outcomes in hospitalized patients with COVID-19: a global literature survey. Critical Care, 24(1), pp.1-4.
5. Abdool Karim, S.S. and de Oliveira, T., 2021. New SARS-CoV-2 variants—clinical, public health, and vaccine implications. New England Journal of Medicine, 384(19), pp.1866-1868.
6. Davies, N.G., Abbott, S., Barnard, R.C., Jarvis, C.I., Kucharski, A.J., Munday, J.D., Pearson, C.A., Russell, T.W., Tully, D.C., Washburne, A.D. and Wenseleers, T., 2021. Estimated transmissibility and impact of SARS-CoV-2 lineage B. 1.1. 7 in England. Science, 372 (6538).
7. Ali, F., Elserafy, M., Alkordi, M.H. and Amin, M., 2020. ACE2 coding variants in different populations and their potential impact on SARS-CoV-2 binding affinity. Biochemistry and biophysics reports, 24, p.100798.
8. Sartorius, N., 2015. Why do we need a diagnosis? Maybe a syndrome is enough?. Dialogues in clinical neuroscience, 17(1), p.6.
9. Tromberg, B.J., Schwetz, T.A., Pérez-Stable, E.J., Hodes, R.J., Wochik, R.P., Bright, R.A., Fleurence, R.L. and Collins, F.S., 2020. Rapid scaling up of Covid-19 diagnostic testing in the United States—the NIH RADx initiative. New England Journal of Medicine, 383(11), pp.1071-1077.
10. Pizzol, J.L.D., Hora, V.P.D., Reis, A.J., Vianna, J., Ramis, I., Groll, A.V. and Silva, P.A.D., 2020. Laboratory diagnosis for Covid-19: A mini-review. Revista da Sociedade Brasileira de Medicina Tropical, 53.
11. La Marca, A., Capuzzo, M., Paglia, T., Roli, L., Trenti, T. and Nelson, S.M., 2020. Testing for SARS-CoV-2 (COVID-19): a systematic review and clinical guide to molecular and serological in-vitro diagnostic assays. Reproductive biomedicine online, 41(3), pp.483-499.
12. Jiang, H.W., Li, Y., Zhang, H.N., Wang, W., Yang, X., Qi, H., Li, H., Men, D., Zhou, J. and Tao, S.C., 2020. SARS-CoV-2 proteome microarray for global profiling of COVID-19 specific IgG and IgM responses. Nature communications, 11(1), pp.1-11.
13. de Abreu, M.C., Choquet, C., Petit, H., Bouzid, D., Damond, F., Marot, S., Ferre, V.M., Burrel, S., Boutolleau, D., Houdou-Fidouh, N. and Marcelin, A.G., 2020. SARS-CoV-2 IGM and IGG rapid serologic test for the diagnosis of COVID-19 in the emergency department. The Journal of Infection, 81(5), p.816.
14. Ghaffari, A., Meurant, R. and Ardakani, A., 2020. COVID-19 serological tests: how well do they actually perform?. Diagnostics, 10 (7), p.453.
15. Bao, L., Deng, W., Gao, H., Xiao, C., Liu, J., Xue, J., Lv, Q., Liu, J., Yu, P., Xu, Y. and Qi, F., 2020. Lack of reinfection in rhesus macaques infected with SARS-CoV-2. BioRxiv.
16. Godlee, F. and Silberner, J., 2020. The BMJ interview: Anthony Fauci on covid-19. bmj, 370.
17. O'Connell, J., 2002. The basics of RT-PCR. In RT-PCR Protocols (pp. 19-25). Humana Press.
18. Tahamtan, A. and Ardebili, A., 2020. Real-time RT-PCR in COVID-19 detection: issues affecting the results. Expert review of molecular diagnostics, 20(5), pp.453-454.
19. Wu, J., Liu, X., Zhou, D., Qiu, G., Dai, M., Yang, Q., Pan, Z., Zhou, N. and Wu, P., 2020. Identification of RT-PCR-negative asymptomatic COVID-19 patients via serological testing. Frontiers in public health, 8, p.267.
20. Whitman, J.D., Hiatt, J., Mowery, C.T., Shy, B.R., Yu, R., Yamamoto, T.N., Rathore, U., Goldgof, G.M., Whitty, C., Woo, J.M. and Gallman, A.E., 2020. Test performance evaluation of SARS- CoV-2 serological assays. Nature biotechnology, 38(10), p.1174.
21. Hachim, A., Kavian, N., Cohen, C.A., Chin, A.W., Chu, D.K.,

DOI: 10.36838/v4i6.16

Mok, C.K., Tsang, O.T., Yeung, Y.C., Perera, R.A., Poon, L.L. and Peiris, J.M., 2020. ORF8 and ORF3b antibodies are accurate sero logical markers of early and late SARS-CoV-2 infection. Nature immunology, 21(10), pp.1293-1301.
22. Cavalera, S., Colitti, B., Rosati, S., Ferrara, G., Bertolotti, L., Nog arol, C., Guiotto, C., Cagnazzo, C., Denina, M., Fagioli, F. and Di Nardo, F., 2021. A multi-target lateral flow immunoassay enabling the specific and sensitive detection of total antibodies to SARS CO V-2. Talanta, 223, p.121737.
23. Chia, P.Y., Ong, S.W.X., Chiew, C.J., Ang, L.W., Chavatte, J.M., Mak, T.M., Cui, L., Kalimuddin, S., Chia, W.N., Tan, C.W. and Chai, L.Y.A., 2021. Virological and serological kinetics of SARS-CoV-2 Delta variant vaccine-breakthrough infections: a multi-cen ter cohort study. MedRXiV.
24. Tahan, S., Parikh, B.A., Droit, L., Wallace, M.A., Burnham, C.A.D. and Wang, D., 2021. SARS-CoV-2 E gene variant alters analytical sensitivity characteristics of viral detection using a com-mercial RT-PCR assay. Journal of Clinical Microbiology, pp.JCM-00075.
25. Mansfield, E.A., 2014. FDA perspective on companion diagnostics : an evolving paradigm.
26. Moran, A., Beavis, K.G., Matushek, S.M., Ciaglia, C., Francois,N., Tesic, V. and Love, N., 2020. Detection of SARS-CoV-2 by use of the Cepheid Xpert Xpress SARS-CoV-2 and Roche cobas SARS-CoV-2 assays. Journal of clinical microbiology, 58(8), pp.e00772-20.
27. Moulahoum, H., Ghorbanizamani, F., Zihnioglu, F., Turhan, K., & Timur, S. (2021). How should diagnostic kits development adapt quickly in COVID 19-like pandemic models? Pros and cons of sensory platforms used in COVID-19 sensing. Talanta, 222, 121534 . https://doi.org/10.1016/j.talanta.2020.121534
28. Pettit, S.D., Jerome, K.R., Rouquié, D., Mari, B., Barbry, P., Kan da, Y., Matsumoto, M., Hester, S., Wehmas, L., Botten, J.W. and Bruce, E.A., 2020. 'All In': a pragmatic framework for COVID-19 testing and action on a global scale. EMBO molecular medicine, 12(6), p.e12634.
29. Xiaojie, S., Yu, L., Guang, Y. and Min, Q., 2021. Neutralizing anti bodies targeting SARS-CoV-2 spike protein. Stem cell research, 50, p.102125.
30. Vandenberg, O., Martiny, D., Rochas, O., van Belkum, A. and Kozlakidis, Z., 2021. Considerations for diagnostic COVID-19 tests. Nature Reviews Microbiology, 19(3), pp.171-183.

Author

Siddhartha Pulukuri is a Senior at Sharon High School who has dedicated his time to researching breast cancer at Boston University School of Medicine since the start of high school. He hopes to major in either Microbiology or Immunology.

IJHSR
ijhighschoolresearch.org
■ REVIEW ARTICLE

The Intricacies of Bioprinting Vascular Structures

Tarun C. Rao
Leland High School, 6677 Camden Ave, San Jose, 95120, CA, USA; taruncrao12@gmail.com

ABSTRACT: The field of tissue engineering and regenerative medicine has seen great advances in the last two decades. The overall implementation of these synthesized structures is still, however, greatly limited. One of the major contributors to this stagnation is the lack of functional vasculature within these structures. Given the role that vasculature plays in the distribution of oxygen and nutrients within the body, *in vitro* organs must also have this network of blood vessels to function properly. Herein, we review state-of-the-art bioprinting techniques and propose the best viable method amongst them for the creation of vascular structures. The three methods we analyzed are extrusion-based, jetting-based, and stereolithography. Information on specific bionics and the overall process for each method is provided, sourced from the primary and secondary scientific literature. While our findings indicated that extrusion-based bioprinting is the most popular bioprinting method used for printing vasculature, accounting for a majority of the papers reviewed it was determined stereolithography bioprinting holds the greatest potential in the creation of functional in vitro vasculature. This decision was made given its high resolution, affordable cost, and high cell viability percentage.

KEYWORDS: Biomedical Engineering; Cell and Tissue Engineering; Vasculature; Bioprinting; Bioink.

■ Introduction

Within the past decade, the field of regenerative medicine has gained recognition as a legitimate option for the regeneration of functional tissue and organs. One of the major contributors to this improved viability is the development of three-dimensional (3D) bioprinting. 3D bioprinting, like 3D printing, is an additive manufacturing technique. A solution containing cells and support material commonly referred to as "bio-ink" is deposited onto a support stage or liquid medium. It is then incubated to form a fully functioning tissue via crosslinking. Crosslinking refers to a bond that links a multitude of polymers. These bonds can be either covalent or ionic bonds. Bioprinting methodologies can be classified into one of two methods: Scaffold or scaffold-free. Scaffold bioprinting involves the printing of cells with support material to support the growth and proliferation of the cells.[1] The scaffolds mimic the extracellular matrix (ECM) as *in vivo*; the ECM provides structural support and can promote regeneration. An added benefit of scaffold bioprinting is the ability to support a higher cell density while also creating more complex geometries, unimpaired by gravity. Alternatively, scaffold-free bioprinting does not include any support material that is printed with the cells. Rather, the cells are first clumped together to form aggregates or clusters. From there, they secrete ECM for structural stability and can then be printed in pellets.[1] With scaffold-free bioprinting, cells are given more room for interaction resulting in the development of a 3D structure that more closely facilitates the high levels of cell communication we see in native tissue.[2] Bioinks, in this process, are limited to just tissue strands and tissue spheroids. Finally, an important aspect of bioprinting is the resolution of the structure. The print resolution refers to "the lowest measurable unit of printed material in the x and y dimensions" (Lee).[3] The lower the unit, the higher the resolution.

While bioprinting has had moderate success in the printing of simple tissue structures like skin, bladders, and ears, the creation of complex large-scale organs such as hearts and kidneys is yet to be seen *in vitro*. As of now, engineered organs are only a fraction of the size of their *in vivo* counterparts. This can be attributed to the lack of movement concerning nutrients within large-scale engineered organs. Nutrients enter tissue through a process called diffusion, where a substance moves from an area of high concentration to an area of low concentration. The rate at which this process occurs is known as the diffusion rate. Although important, the diffusion distance plays an equal, if not larger, role in the usefulness of nutrients. The larger an object is, the larger the diffusion distance. Ultimately, resources for cells could be used up before even reaching their destination or crossing a barrier, no matter how efficient the diffusion rate is. If a structure is too large to where it cannot receive nutrients due to a large diffusion distance the cells will die, hence rendering the organ nonfunctional. Instead of this slow and inefficient process, the use of vasculature will transport nutrients to where they are required in a much timelier manner. For example, instead of relying on sugar to diffuse through a multitude of tissues to go from the stomach to the muscles, vasculature facilitates the delivery of nutrients and oxygen throughout the tissue, preventing mass necrosis or cell death throughout the structure. Ultimately, this is the main challenge that is preventing the implementation of synthesized organs. In this review, we examine the structure of vasculature, the composition, and properties of bioinks, and novel bioprinting methods, assessing which bio-

DOI: 10.36838/v4i6.17

printing method has the best chance of achieving the mass generation of a functional vasculature.

Discussion

The vascular system:

The vascular system remains one of the important organ systems within our bodies. The main function of the system is the transportation of nutrients such as sugar, oxygen, blood, and hormones towards and away from different tissues and organs using different types of blood vessels. An example of this is the transport of insulin through the bloodstream. In tissue engineering and bioprinting, a functional vascular system is of the utmost importance concerning the survival of the synthesized tissue.

Blood Vessels Composition and Function:

There are three types of blood vessels: arteries, veins, and capillaries. The job of these vessels involves the transportation of oxygenated and deoxygenated blood to and from tissues. There are two systems of blood vessels: the pulmonary vessels and the systemic vessels. The pulmonary vessels transport blood from the right ventricle of the heart to the lungs and back to the left atrium. In contrast, the systemic vessels transport blood from the left ventricle to all the tissue in the organism, returning the blood to the right atrium.

Arteries Composition and Function:

Arteries are the blood vessels responsible for carrying blood away from the heart and the function is divided into pulmonary and systemic arteries. Pulmonary arteries transport deoxygenated blood from the right ventricle to the lungs. There, the blood is oxygenated by the lungs. Contrary to the pulmonary arteries, the systemic arteries transport oxygenated blood from the left ventricle to all tissue within the body. The further the blood is transported from the heart, the smaller the arteries become, branching out like a tree. The smallest arteries are known as arterioles.

Arteries are composed of three layers of tissue. These layers are the tunica adventitia, also known as the tunica externa, the tunica media, and the tunica intima as shown in Figure 1. The tunica adventitia is the outermost and thickest layer of the artery. It is composed of collagen and elastic tissues and the vasa vasorum. The vasa vasorum is a network of tiny blood vessels that can supply blood to larger blood vessels. They are the vessels of vessels. The main function of the tunica adventitia is to prevent the vessel from overextending and bursting. The middle layer of the artery, the tunica media, is composed of smooth muscle cells, elastic tissues, and collagen. In addition to providing support for the artery, the tunica media is also responsible for changing the vessel's diameter, increasing, or decreasing the blood flow and blood pressure within the vessels. The tunica intima is the innermost layer of the artery and is composed solely of endothelial cells. Endothelial cells are a specific type of cell that forms a boundary between a vessel and tissue and are responsible for regulating the distribution of substances that enter and exit tissue. Additionally, certain signals released by endothelial cells can induce the growth of vessels.[4]

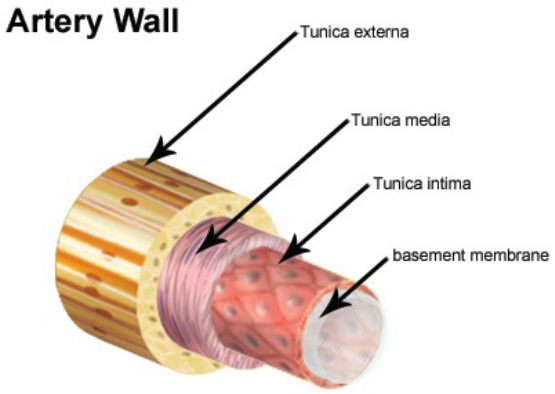


Figure 1: The layers of tissue in an artery are shown. Image referenced from National Cancer Institute SEER Training.[5]

Veins Composition and Function:

In contrast to the function of arteries, veins are responsible for carrying blood back from the tissues toward the heart. Most of the veins within the body carry deoxygenated blood back to the heart after the exchange of nutrients has occurred in tissues. However, there are exceptions to this function such as pulmonary veins. Both veins carry oxygenated blood back to the heart. Veins are composed of the same three layers of tissue as arteries: The tunica externa, tunica media, and tunica intima. However, there is less smooth muscle and tissue and thus the walls of veins are thinner compared to arteries.[5] Some unique structural features that veins have specifically are valves to prevent backflow. During transportation, blood is transported to the right ventricle to be pumped to the lungs to be oxygenated again. With pulmonary veins, oxygenated blood is returned from the lungs to the left atrium and is ejected into the left ventricle.

Capillaries Composition and Function:

The third type of blood vessel within the vascular system is a capillary. While veins and arteries are responsible for the movement of blood throughout the body, capillaries are responsible for the transfer of oxygen, nutrients, and waste between tissue and blood. This transfer occurs through passive diffusion, where a substance travels down a concentration gradient through a membrane. Additionally, this process can also occur through pinocytosis, where vesicles (membrane-bound sacs) take in nutrients and merge with the cellular membrane. Unlike veins and arteries, capillaries are composed of only two layers of tissues. There is an inner layer composed of endothelial cells and an outer layer composed of epithelial cells.

There are three primary classifications of capillaries: continuous, fenestrated, and sinusoidal. Continuous capillaries are, as the name suggests, continuous and uninterrupted, allowing only small molecules to pass. Fenestrated capillaries have small pores that allow the passage of slightly large molecules. These capillaries are found in areas that frequently experience exchanging of blood such as the small intestine and kidneys. Sinusoidal capillaries are the most disrupted capillary type as the large pores within the capillary allow for the passage of large molecules and even cells. Sinusoidal capillaries are found in tissues such as the liver, spleen, and bone marrow.

Growth of Blood Vessels:

The growth of new blood vessels is vital to the function and health of an organism. Subsequently, the independent growth of vasculature after fabrication will give a synthesized organ much higher viability. Two mechanisms form blood vessels: vasculogenesis and angiogenesis.

Vasculogenesis is the process of creating blood vessels in the embryo. It is the first stage in the development of the circulatory system and is often followed by angiogenesis. One key aspect that separates vasculogenesis from angiogenesis is that vasculogenesis occurs where there are no pre-existing blood vessels. Vasculogenesis occurs through the differentiation of endothelial precursor cells known as angioblasts and through that differentiation, a de novo formation of a vascular network is created.[6]

Unlike vasculogenesis, angiogenesis is the formation of new blood vessels from pre-existing vessels. The initiation of angiogenesis is controlled by several growth factors such as vascular endothelial growth factor (VEGF), fibroblast growth factor (FGF), and tumor necrosis factor (TNF-α).[7] A growth factor is a protein that can instruct a specific cellular response. VEGF is responsible for promoting the growth of new proteins and forms a part of the mechanism that restores the blood supply when there is compromised blood circulation. There are two types of angiogenesis: Sprouting angiogenesis and intussusceptive angiogenesis.

Sprouting angiogenesis (SA) is characterized by the growth of endothelial cells and this type of angiogenesis can add blood vessels to areas that did not have blood vessels previously. It is initiated by the detection of hypoxia or a lack of oxygen. In these areas parenchymal cells (functional cells of organs, e.g., neurons) secrete VEGF. Subsequently, tip cells travel toward the growth factor, leading to the growth of new capillary sprouts. Concurrently, endothelial cells become stalk cells and build the blood vessels behind the tip cells.[8] In simpler terms, SA can be thought of as the creation of a deviation from the main blood vessel such as an exit lane from the main highway. As shown in Figure 2, a new deviation from the main vessel is being created as the tip cell leads the development of the new vessel.

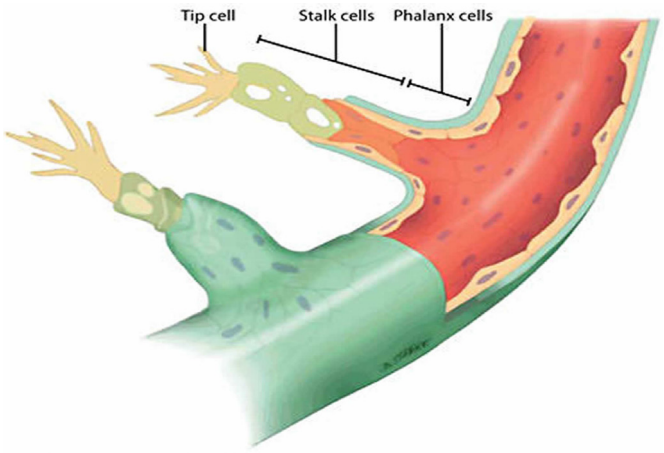


Figure 2: The sprouting angiogenesis process is shown. Image referenced from Spiegelaere *et al.*[9]

The discovery of intussusceptive angiogenesis occurred in the 1980s and is a novel and relatively unexplored phenomenon. Intussusceptive angiogenesis (IA) can be thought of as the splitting of a pre-existing blood vessel into two smaller copies, somewhat similar to DNA replication. A defining characteristic of IA is the formation of intraluminal tissue pillars, formed when endothelial walls of the vessel migrate towards each other, similar to cytokinesis in mitosis and meiosis. As shown in Figure 3 section C, after the initial pillar is created, pericytes and myofibroblasts (vascular mural cells in the basement membrane),[10] inject an extracellular matrix into the pillar and finally, two new vessels are formed.[9] IA can only occur in a network formed through vasculogenic or SA. Typically, the initial network is formed through SA and is further developed by IA.

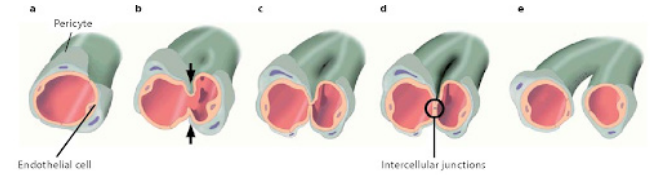


Figure 3: The Process of intussusceptive angiogenesis is shown. Image referenced from Spiegelaere *et al.*[9]

Extrusion-based bioprinting:

Given the uses of vasculature within an organism, it has been made evident that the need for functional blood vessels is dire. Luckily, bioprinting is slowly removing this roadblock, opening the door for functional, life-size, and implantable *in vitro* organs.

One of the most common printing methods is extrusion-based bioprinting due to its low production cost and wide range of printable biomaterials. The printable viscosities of biomaterials that can be extruded range from 30 millipascal seconds (mPa/s) to more than 6 * 107 mPa/s.[11] This covers a range of materials that have the viscosity of motor oil up to window putty including, honey, lard, and ketchup. In extrusion-based bioprinting, a continuous filament of a cell-laden ink is extruded out of a nozzle by a pneumatic system or a screw. Layers of bioink are serially built on top of one another to form a three-dimensional figure following a CAD model. After the primary structure is formed via extrusion, ultraviolet (UV) light is used to induce crosslinking between polymers, forming a durable and viable structure (Figure 4).

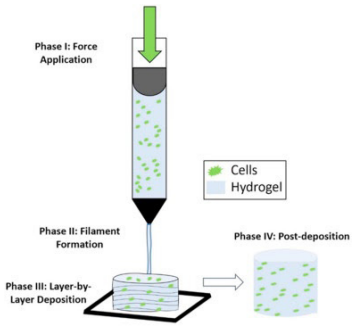


Figure 4: The various phases of Extrusion-Based Bioprinting (EBB) are shown. Phase 1: The bio-ink consists of various cells and a hydrogel medium for stability and force is applied to force the bio-ink out. Phase 2: The bio-ink is extruded in a continuous filament. Phase 3: The bio-ink is deposited in

DOI: 10.36838/v4i6.17

layers, forming a three-dimensional structure. Phase 4: Post-extrusion modifications are shown, which include the crosslinking of hydrogels. This fuses the layers, ensuring mechanical durability. Image referenced from Ramesh *et al.*[12]

Extrusion Printing Methods:

There are three categories of methods by which bioink can be extruded out of the nozzle. These methods are pneumatic, screw-based, and piston-based (Figure 5). In pneumatic bioprinting, the bio-ink is forced out of the nozzle using pressurized air. Contrary to a pneumatic system a screw-driven printer uses a rotating screw, driven by a motor, to extrude the bio-ink.[13] The benefits of using a pneumatic system include its simplicity and lack of maintenance needed, yet accuracy may differ based on the viscosity of the bio-ink. Screw-driven methods provide more direct control over the placement of the bio-ink. However, screw-driven extrusion results in a higher percentage of cell damage in comparison to pneumatic extrusion bioprinting.[13] This cell damage can be attributed to the shear stress that cells experience when being extruded as the depth of the screw threads and geometry of the screw may deform the cells themselves. Ning *et al.*[13] found that screws with a smaller pitch distance and higher thread depth are more likely to cause cell damage. The use of a piston in extrusion-based bioprinting is very similar to pneumatic-based bioprinting where a piston is used instead of pressurized air.

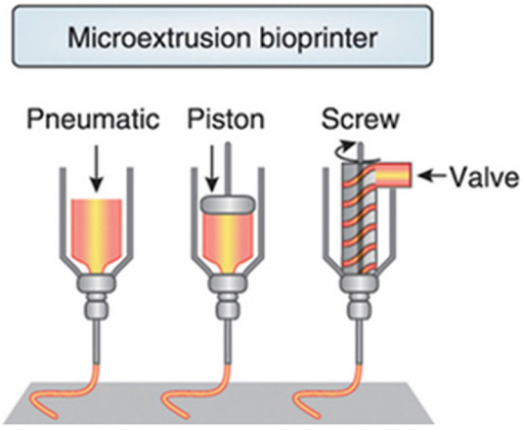


Figure 5: This image has been referenced from Wilson *et al.*[14] The three different methods of extrusion-based bioprinting are shown.

Bio inks:

The variety of bioinks that can be utilized in extrusion-based bioprinting is immense. Several biomaterials are compatible with the printing method including hydrogels[14-18], spheroids[19-20], micro-carriers, and the extracellular matrix.[21] Due to the flexibility of nozzle diameters, bioinks of many viscosities can be extruded.

Hydrogels:

A hydrogel is a three-dimensional substance that is composed of hydrophilic polymers and can retain large amounts of water while maintaining its shape. Hydrogels are often used as bioinks as they mimic the extracellular matrix and are biocompatible while also having high printability.[22]

There are three categories of hydrogels: natural, synthetic, and hybrid. Natural hydrogels can be classified as hydrogels whose polymers originate from organic material like collagen, and other components of the ECM.[23] As a result, an environment that enhances cellular proliferation is acquired, but at the cost of durability as natural hydrogels have a weak structure and the distortion of one layer could subsequently destroy the printed tissue.[24] Compared to natural hydrogels, synthetic hydrogels are created using synthetic polymers such as polyamides. Since synthetic hydrogels are, as the word suggests, "synthetic" they lack the big interactive capabilities of natural hydrogels.[25] However, this disadvantage is somewhat offset by their mechanical strength and durability. As the name suggests hybrid hydrogels are a combination of both synthetic and natural hydrogels. They are composed of proteins, peptides, and can even contain nano/microstructures, and are interconnected, undergoing chemical modifications.[26] Thus, they retain the bioactivity seen in natural hydrogels as well as some of the mechanical properties visible in synthetic hydrogels. Some common properties that all hydrogels share are swelling in which they can take in water like a sponge while retaining their physical orientation.

Alginate is an example of a natural hydrogel. It is a hydrophilic polymer that originates from the cell wall of seaweed and brown algae. It is composed of d-mannuronic acid and a-L-guluronic acid.[27] The carboxylic chains within the structure allow for ionic crosslinking that is often induced by a CaCl2.[18] In 3D bioprinting, alginate remains one of the most widespread biomaterials due to its biocompatibility and ability to form porous structures.

Gelatin is an example of a hybrid hydrogel that is the result of structural degradation of collagen and the process of gelation occurs when polymer chains either undergo physical or chemical crosslinking. Thermal gelation is one method of gelation that forms a gel by heating a solution, inducing cross-linking. The chemical modification of gelation with methacrylic anhydride forms a photocrosslinkable natural bioink and the Gelatin methacryloyl (GelMA).[17] Hence it can be classified as a hybrid hydrogel. GelMA can be categorized into two types, one that is produced from acid treatment (Type 1) and the other produced from alkali treatment (Type 2). In contrast to pure gelatin, which is soluble at body temperature, GelMA can maintain its structural form at body temperature.[17] This makes GelMA a much more viable bioink as the 3D structure formed will not be degraded by the heat of the human body when implanted in vivo. Additionally, GelMA is shown to have unparalleled scaffold strength, enabling it to support cell viability for 14 days.[17] A study conducted by Lee *et al.*[17] found that both Type A and Type B constructs had high cell viability when the cell concentration was at 20%, indicating that GelMA is a viable bioink for extrusion-based bioprinting.

PluronicF127 is an example of a synthetic hydrogel as it is composed of amphiphilic copolymers that contain ethylene oxide and polypropylene oxide.[28] While it is a synthetic hydrogel it has some natural hydrogel properties such as biocompatibility and biodegradability. A unique characteristic of PluronicF127 is thermosensitivity which can cause a sol-gel phase transition. This phase can be defined as the change from a liquid to a gel. Due to this phase PluronicF127 can encapsulate cells better and be more adhesive.[28] PluronicF127 has been applied in the

creation of vasculature as Suntornnond *et al*.[15] used a combination of PluronicF127 and GelMA for a bioink.

Spheroids:

Recently the use of tissue spheroids as potential bioink has gained more popularity as the benefit of using tissue spheroids is that instead of having cells attempt to proliferate through a hydrogel, the starting number of cells can be relatively high which will then cause ECM to be deposited, eventually forming a 3D structure.[29]

Tissue spheroids refer to three-dimensional cell aggregates that are in a spherical orientation. Cell aggregates refer to cells that have clustered together. To date, the primary usage of these spheroids has been in cancer research and the discovery of new drugs as the testing of novel drugs can be done *in vitro*.

Several methods are applicable in the printing of tissue spheroids. As always, however, certain restrictions and considerations must be taken care of before printing. The two most important requirements concerning tissue spheroids are that during the extrusion process, the pressure and shear stress placed on the spheroids should not cause major DNA or structural damage and the ability of fusion is not compromised.[19] The process of fusion is shown in Figure 6 and one can see that the fusion of the spheroids is integral to the development of a smooth and interconnected structure.

To use tissue spheroids as a bio-ink, one must first generate the spheroids themselves. Additionally, the spheroids themselves must be uniform as their use in organ bioprinting is largely dependent on their scalability. In 2011, a novel method of tissue bio fabrication was developed by Mehesz *et al*.[30] Some previous methods that were used were centrifugation, gravity-reliant hanging drops, and cultivation within a flask. However, all these methods have size and shape parameters outside of the experiment's control and thus their scalability is unreliable. Mehesez *et al*. utilized wells that were filled with non-adhesive hydrogels. From there, suspended cells would form spheroids, with their weight dragging them towards the rounded bottom, forming a spherical shape. To dispense the suspended cells a pipette system was used. Tissue spheroids were seen just 48 hours after being placed in the micro recessions.

In the past decade, success has been found in the printing of tissue spheroids in 2019 Aguilar *et al*.[20] and was used to print mesenchymal stem cells using a Regenova printer. The target diameter for the printer was 500 μm. To generate the spheroids, Aguilar et al. used centrifugation and gravity and cultured each group for up to 28 days. After printing the viability of all cells present excluding a cell group known as trypan blue was higher than 85%. Additionally, it was seen that the spheroids cultured in an osteogenic medium expressed a higher level of osteogenic genes and the DNA content of all cells increased each day. These findings give light to the fact that the use of tissue spheroids as bioink can increase proliferation and cellular interactions, and generate a viable 3D structure that truly mimics an *in vivo* organ.

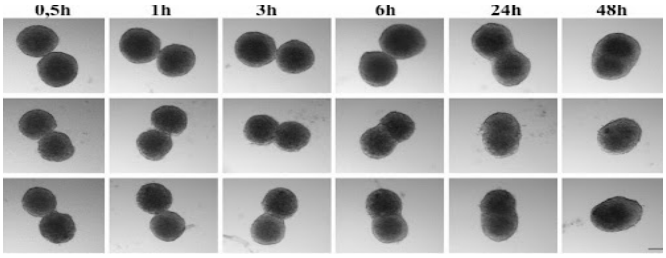


Figure 6: The process of fusion is shown by three samples of tissue spheroids over 48 hours. The image is referenced from Kosheleva *et al*.[31]

Micro-carriers:

Micro-carriers (MC) are support matrices that help cells grow in bioreactors. They are usually spherical with diameters of 150-200 μm.[32] Their development in bioprinting has stemmed from certain limitations that occur within the use of hydrogels in bioprinting. One problem that may occur with hydrogels is that no matter the *in vivo* cellular environment, in the bio-ink, the cell is placed in a hydrophilic, or water-loving environment and is constrained to a spherical shape.[33] A benefit of using MCs is that they can be embedded in hydrogels and can increase the mechanical strength of the hydrogel in the process.[33]

Decellularized Extracellular matrix components:

While biomaterials such as hydrogels and micro-carriers do provide the cells with a safe and durable environment, a major flaw that they possess is that they cannot accurately replicate the behavior and function of the extracellular matrix (ECM). This can be attributed to the fact that many hydrogels are composed of a singular component of the ECM such as collagen, instead of the ECM in its entirety. The purpose of the ECM is to regulate the cell shape as well as initiate many intracellular reactions and is a vital component in cell communication. Additionally, the ECM contains many growth factors which are key to the development of functional tissue. In terms of bioprinting, an environment that has a resemblance to the ECM is integral to fabrication as after the primary structure is printed, proliferation and differentiation are needed to mimic an *in vivo* organ or tissue structure. A simple solution to this problem is the use of decellularized ECM as a bio-ink. For clarification, decellularized ECM (dECM) simply refers to ECM that has been isolated from its cellular environment. To create a dECM bioink, first, the ECM is removed from cells, and then solubilized, and finally, the pH is adjusted to prevent cell damage due to the acidic pH of the ECM.[21] A study conducted by Pati *et al*. found that the use of dECM bioink resulted in the formation of ECM within the 3D printed structure, high cell viability, and no stress-induced apoptosis during extrusion.[21]

Disadvantages:

While extrusion-based bioprinting remains one of, if not the most, accessible bioprinting methodologies it has several key flaws that have spurred on the development of drop-based bioprinting and stereolithography. One particular disadvantage that occurs in extrusion-based bioprinting is the tendency for cells to be damaged due to shear stress, impacting the functionality and viability of the cells (as low as 40% viability).[34] This occurs during the extrusion of bio-ink through a nozzle of a small diameter and can also occur when the bioink is being

DOI: 10.36838/v4i6.17

extruded by a screw-driven process.[13] In terms of nozzle diameter, a nozzle with a smaller diameter will result in greater accuracy but it comes at the cost of cell viability as the smaller diameter results in greater stress which will deform the cells due to the lack of space. For clarification, shear stress is when a force acts parallel to the area of a cross-section, unlike normal force which acts perpendicular to the area. Examples of shear stress include chewing food, applying the brake in a car, or running. While the issue of shear stress does play an impact on cell viability the scope of the issue has been reduced due to the introduction of shear-thinning bioinks such as alginate. Shear-thinning refers to how the viscosity of the gel decreases with an increase in stress, mitigating the amount of stress that affects the ink. This can result in higher cell viability as the amount of deformation within the bio-ink will decrease.

Applications

FRESH Bioprinting:

A common problem that occurs with many bioinks in extrusion-based bioprinting is that they are soft. Unlike plastics and metals used for 3D printing, biomaterials such as collagen, alginate, decellularized ECM, and many others do not have the durability to maintain their geometry after being extruded. This phenomenon often occurs when these materials are extruded onto a flat support stage, such as a petri dish. However, researchers at Carnegie Mellon University, directed by Professor Adam Feinberg have developed a bioprinting method that involves the "freeform reversible embedding of suspended hydrogels", otherwise known as FRESH The purpose of FRESH is to embed the printed hydrogels within a secondary hydrogel material composed of gelatin microparticles. However, one key innovation within the hydrogel bath is that the material acts as a high viscous material when exhibiting low shear stress but acts as a low viscous material when exhibiting high shear stress.[35] In layman's terms, the extrusion needle will encounter little resistance when penetrating the hydrogel and when exiting, yet when the bioink is extruded it will keep its current orientation as the secondary hydrogel will exhibit a high viscosity. After the bioink achieves crosslinking, forming a viable 3D structure, the support bath can then be melted in a non-destructive manner by raising the temperature. Hinton *et al.*[35] were able to successfully print a viable femur and heart; structures that would otherwise be impossible to print given the durability of these soft hydrogels. Thus, using FRESH the orientation of bio-inks will no longer be constrained by the weight or durability while still in gel form.

Creating Vasculature:

As of today, there has been a multitude of studies that have successfully developed perfusable vasculature. Suntornnond *et al.*[15] were able to use a composition of Pluronic F127 and GelMA to create a hydrogel and used human umbilical vein endothelial cells (HUVEC) to create a vascular structure. After printing, the cells were cultured for 7 days and achieved a maximum number of 600,000 cells, staying alive until day 10.

To create their vasculature conduits Zhang *et al.*[36] used human umbilical vein smooth muscle cells (HUVSMCs) and alginate as the bio-ink. A coaxial extrusion system was used to create the conduits. The initial viability of the cells was 73±2%

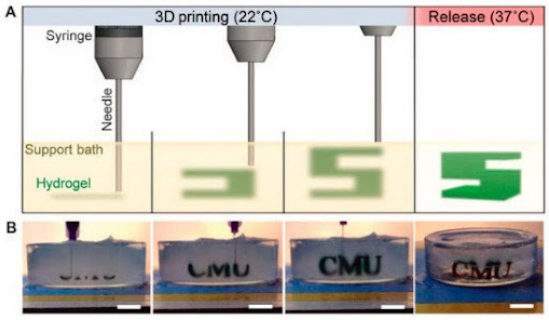


Figure 7: A visual representation of the FRESH technique can be seen. (A). A needle containing a hydrogel-based ink extrudes said ink into a support bath at a temperature of 22 degrees Celsius. Layer by layer the figure is built. After the figure has been oriented and cross-linked, the temperature is raised to 37 degrees Celsius, melting the support material, leaving only the finished structure. (B). An example is shown with the letters "CMU" being printed. This image was used from Hinton *et al.*[35]

and on day 7 the cells had maximum viability of 84±1%. Additionally, the conduits also showed high perfusion with 405±11 µl per hour. Finally, extracellular matrix development was seen on the printed cells, indicating that the cells are communicating and becoming more like their *in vivo* counterparts.

Like Zhang *et al.* Tabriz *et al.*[18] also used alginate as the support material for the bio-ink with Human glioma U87-MG cells being the cell culture. An interesting procedure conducted by Tabriz and Co. was that they partially crosslinked their alginate with CaCl2 and after extrusion, the cell-laden alginate was again crosslinked. Finally, after the second crosslinking BaCl2 was used to perform one final crosslinking. After printing the cells had a viability of 92.9 ±+0.9% and over 11 days the cell viability was always above 82%.

Another study that was able to create vascular structures was Jia *et al.*[37], which used a blend bioink composed of GelMa, alginate, and 4 arm polyethylene glycol-tetra-acrylate (PEG-TA). Like Tabriz *et al.* they partially crosslinked the bioink to form stable constructs. The printing system used multilayered coaxial nozzles with HUVECs and human mesenchymal stem cells (MSCs) being the cell culture. After printing the bio-ink, the constructs were exposed to UV light to induce covalent crosslinking. It was found that the cell viability for UV exposure times of 20 seconds and 30 seconds was above 80% after 7 days of culture.

Extrusion-based bioprinting has great potential in the field of tissue engineering and regenerative medicine. This can be attributed to the versatility of bio-inks as well as the low cost of synthesis. Additionally, there is a great variety in printers as the website Aniwaa listed 11 bioprinters.[38] Out of those 11, 6 printers could perform extrusion-based bioprinting, all for affordable prices. The highest-priced was the BioScaffold Printer BS3.2, priced at $150,000. In stark contrast the lowest-priced bioprinter, Dr. INVIVO 4D was priced at $20,000. Products of extrusion-based bioprinting can be used in tissue transplants as well as seeing the effects of certain drugs on cells with zero risks to a patient's life. This could result in increased development of experimental drugs as well as personalized organ transplants, eliminating the need for donor lists. These studies do indicate that the development of functional vasculature is plausible using extrusion-based bioprinting.

Droplet-Based Bioprinting:

Droplet-based bioprinting is an additive manufacturing technique where a cell-laden bioink is extruded out of a nozzle in the form of droplets. There are three overarching methods of droplet-based bioprinting, each with its sub-methods. These three methods are inkjet bioprinting, acoustic droplet ejection, and micro-valve bioprinting. Inkjet bioprinting has several sub-methods but the two most common methods are drop-on-demand bioprinting and electrohydrodynamic bioprinting. Finally, drop-on-demand has three sub-methods that each involve different actuators, specifically thermal and piezoelectric actuators, and electrostatic forces. The general process behind droplet-based bioprinting is very similar to extrusion-based bioprinting where droplets are deposited onto a support stage. From there the droplets spread, fusing. Finally, a crosslinking agent is applied to the layer hardening it. The crosslinking agents can vary, consisting of a UV light, a reservoir, mist, or other droplets as shown in Figure 8.

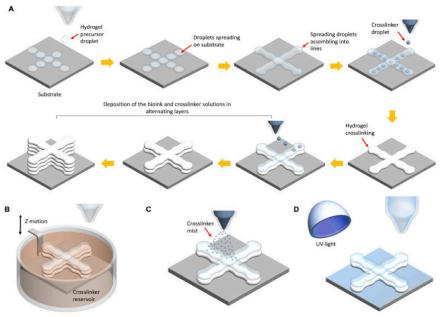


Figure 8: General droplet-based bioprinting process shown with varying crosslinking agents. (A) Bio ink droplets deposited and fused with crosslinking droplets. Multiple layers are deposited (B) Layer fused with a UV light. (C) Crosslinking mist is applied to layers. (D) Layers are placed in a crosslinking solution. Image referenced from Gudapati *et al.*[39]

Acoustic Droplet Ejection:

One major flaw that occurs within nozzle-based bioprinting methods is that the pressure of forcing bioink through a nozzle places mechanical stress on cells, causing cell death, thus reducing the overall viability of the present structure. One method that avoids these disadvantages is acoustic droplet ejection (ADE). In ADE the bioink is placed in a reservoir and acoustic radiation is used to apply force to the bio-ink, generating droplets. Then the droplet is attached to a build platform that is placed above the bioink reservoir. In ADE the size of the droplet is inversely proportional to the frequency of the signal. A higher signal will equate to a droplet with a smaller diameter and vice versa.[40] While this method does have the safety of the cells placed at the forefront it is still a relatively unexplored method, having only been used to create 2D layers and has rarely been used for 3D structures.[40]

Micro-valve bioprinting:

Micro-valve is a bioprinting method that uses an electro-mechanical valve to generate droplets. As seen in Figure 11, micro-valve bioprinting involves the use of pneumatic pressure, a solenoid coil, and a valve. Pneumatic pressure is applied to the bioink using a gas regulator. A voltage pulse generates a magnetic field in a solenoid coil that opens the nozzle and

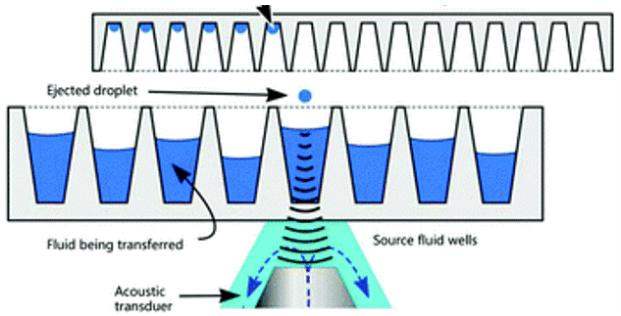


Figure 9: A schematic of the ADE printing process. The transducer (cylinder at the bottom) transmits an acoustic signal that travels upwards. The force generated from the sound generates a droplet that then travels upwards, sticking to the build platform shown at the top of the image. Image referenced from Guo *et al.*[41]

the bioink is then deposited.[41] Two unique factors controlling the deposition are the pressure generated by the gas regulator and the valve opening time.[41] Unlike other inkjet-based methods, microvalve bioprinters often contain multiple printheads, deposit bioinks in synchronization, and print close to 1000 droplets per second with 1-2 μm material thickness.41 Additionally, with a concentration of 1 mil cells/ml Ng and Yeong *et al.*[42] were able to achieve cell viability higher than 95%. However, this result should not be taken at face value since the purpose of the experiment was to determine the minimal cell concentration needed to achieve cell viability above 95%. Rather, in experiments not actively searching for high cell viability, it can be expected that the cell viability will be closer to 86%.[41]

Inkjet Bioprinting and its Sub-Methods:

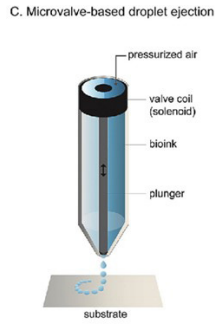


Figure 10: A schematic of a microvalve printhead is shown. Image referenced from Włodarczyk-Biegun *et al.*[43]

Inkjet bioprinting uses gravity, atmospheric pressure, and the fluid mechanics of the bio-ink such as the viscosity to eject droplets.[39]

Electrohydrodynamic Bioprinting:

Electrohydrodynamic (EHD) bioprinters function by using an electric field to pull the bio-ink through a nozzle and EHD is a subset of Inkjet Bioprinting. This benefits the cell viability as the bioink will not experience abnormally large pressure, preventing cell death. As a result of using electricity, the nozzle diameter of nozzles used in electrohydrodynamic printing has to be minuscule (≤ 100 μm) and the bioinks have to be highly concentrated.[39] The process of EHD is that exposure to an electric field causes mobile ions within a liquid to accumulate at the liquid surface. Eventually, a conical shape known as a

DOI: 10.36838/v4i6.17

Taylor cone is created and after the electric field exceeds its critical limit, the stress creates a droplet of bio-ink.[44]

Drop-on-Demand Bioprinting and its Sub-Methods:

Unlike electrohydrodynamic bioprinting, drop-on-demand bioprinting generates droplets of bioink by propelling the bio-ink through the nozzle. This is done by one of three mechanisms: piezoelectric actuators, thermal actuators, or electrostatic forces. For clarification, an actuator is any component that enables movement within a machine. Thermal actuators will simply convert thermal energy into motion. Piezoelectricity refers to the electricity that can accumulate within certain materials such as crystals, DNA, and proteins that are in response to mechanical stress. In simpler terms piezoelectricity refers to electricity that is the result of pressure and latent heat. Thus, piezoelectric actuators take the electric energy that can be generated from the bioink itself and can eject the bio-ink.

Piezoelectric Bioprinting:

Like EHD bioprinting, piezoelectric bioprinting utilizes electricity to generate droplets. Unlike EHD, piezoelectric actuators require crystals such as quartz to conduct electricity. To generate an electric current, the crystal is placed between two metal plates and once mechanical pressure is applied, a current can be forced out of the crystal. In addition to this method, electricity can also be applied to the crystal. This deforms the structure of the crystal releasing energy in the form of a sound wave.[45] In terms of generating a droplet, once the crystal is deformed, the fluid chamber is deformed as well,[39] resulting in a sudden change of volume. The result is that sound waves provide the needed pressure for the bio-ink to overcome the surface tension of the nozzle, delivering a droplet.[46]

Thermal Bioprinting:

Thermal bioprinting is somewhat like piezoelectric bioprinting as both processes require actuators. However, the overall process for thermal bioprinting is significantly more simple than piezoelectric bioprinting. In short, when voltage is applied, the thermal actuator heats the bio-ink. The heating of the bioink generates a bubble and once that bubble pops, pressure is generated which forces out a droplet of bio-ink. A particular concern of thermal inkjet bioprinting is that when the bioink is being heated up the heat will denature proteins and potentially kill cells.

Electrostatic Bioprinting:

Electrostatic bioprinters, like piezoelectric bioprinters, generate droplets by increasing the volume of the fluid chamber. Electrostatic forces refer to forces that can pull or push without physical contact. Unlike thermal and piezoelectric printers, electrostatic printers can eject highly viscous ink.[47] This separates it from other inkjet methods, which can only print low viscosity materials, making it a highly preferred method, given its high-resolution results.

Bio inks:

Due to the nature by which the bioink is deposited in droplet-based bioprinting the viscosity range of bioinks remains quite limited as the use of high viscosity bioinks can result in the lack of droplets. The viscosity range for droplet-based bioinks is 3.5-12 mPa/s.[48] Thus, droplet-based bioprinting is limited to low viscosity bioinks. Within this range, hydrogels remain a popular choice for bioinks. Specific substances include alginate,[49-51] fibrin,[52] collagen,[49] and gelatin.[47] The reason that droplet-based bioprinting requires such low viscosity bioinks is that if the bioinks are in high viscosity there is the potential hazard of the nozzles becoming clogged.

Disadvantages:

Despite its great advantages such as method variety, non-contact nature, and accessibility, droplet-based bioprinting has flaws that prevent it from being a perfect bioprinting method. One of the more prominent flaws is the restriction that is placed on the bioink variety as droplet-based bioprinting has a limited selection of only low viscosity bioinks, else the nozzles will end up clogging up. One theoretical disadvantage present in droplet-based bioprinting and specifically thermal bioprinting is the notion that the high heat (up to 300 degrees Celsius) of the nozzle can denature cells, decreasing cell viability during the printing process.[34] However, Cui *et al*.[53] found this belief to be exaggerated. Using Chinese hamster ovary cells, the reported cell viability was 89% and the temperature of the bioink only increased between 4 and 10 degrees Celsius. Thus, while the limitation of the bioinks can be limiting in the application of droplet-based bioprinting there are little to no more disadvantages besides the bio-inks, making it a very usable method.

Applications:

Despite the limitations, droplet-based bioprinting has been at the forefront of vasculature fabrication. Both Cui and Boland *et al*.[52] as well Yao *et al*.[49] were able to successfully generate vasculature using droplet-based bioprinting. Cui and Boland used fibrin as the bioink with human microvascular endothelial cells as the cell culture. Using thermal inkjet bioprinting they were able to achieve cell viability of 90%. While pores did develop during the deposition of droplets, it was observed that these pores were repaired 2 hours after being deposited. Finally, the fabricated microvasculature remained stable and retained its integrity for 14 days. While also generating microvasculature, Yao et al. used both different cell cultures and bioink. Instead of using human microvascular endothelial cells and fibrin Yao and Co. use HUVECs and alginate microspheres coated in collagen. Additionally, instead of using thermal inkjet bioprinting, Yao and Co. used an electrohydrodynamic printer to dispense the droplets. The generated vasculature was then inserted into mice and the synthetic vascular channels were able to fuse with hose vessels, undergoing vasculogenesis. These two studies do indicate that droplet-based bioprinting is a viable method for generating functional vasculature structures. Finally, droplet-based bioprinting has proven itself to be an easily accessible method as according to Bishop *et al*.[34] droplet-based bioprinting has the lowest cost between itself, extrusion-based bioprinting, and stereolithography. Some printers that can perform droplet-based bioprinting are the BIO X printer and the 3DDiscovery printer, priced at $40,000 and $200,000.

Stereolithography:

Recently, stereolithography apparatus (SLA) has been acknowledged as a bioprinting technique. While extrusion and drop-based bioprinting rely on a nozzle to precisely dispense droplets of cell-laden bioink, SLA uses a light source to selectively illuminate sections of a liquid bioink. The exposure to

the light source induces cross-linking within the vat of bioink, creating a 3D structure.[54]

Process and Components:

When generating a structure using SLA there is a certain amount of variation that can occur when it comes to the structural design of the printer. The methodology used in SLA is that a concentrated light source will be reflected off a mirror into a certain section of photocurable resin, hardening it. However, the position of this light source and mirror can vary as they can either be placed below the resin tank or above (Figure 11).[55] When placed above the resin tank, the light will harden the very top layer of the resin. From there the build platform lowers the completed layers as a recoater blade applies a fresh coat of resin to be hardened. This continues until the structure is completed. In contrast, when the light is underneath the vat of resin, the build platform will immerse itself into the resin and from there the process occurs but in reverse as the build platform will go up as more layers are hardened. When it comes to the light source used for the laser the most common option is to use ultraviolet (UV) light to harden the bio-ink. However, the use of UV light will induce cell damage.[56] To counter this, Wang *et al.*[57] were able to use visible light as the laser and achieved 85% cell viability for at least 5 days.

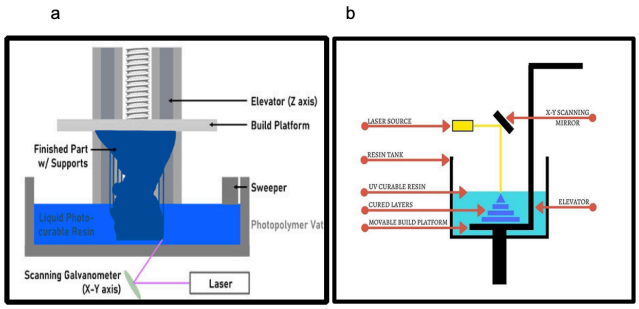


Figure 11: Two variations of stereolithographic printers are shown. (a) A printer with a laser beneath the vat. (b) A printer with a laser above the vat. Image (a) referenced from "All About Stereolithography 3D Printing".[58] Image (b) referenced from All3DP *et al.*[39]

Bio inks:

The bioinks used for SLA are generally derived from bioinks used for both extrusion and droplet-based bioprinting. However, to be applied in SLA these bioinks have to be modified with a functional group to allow for photo-crosslinking. A benefit of photopolymerization is that it can control the formation and structure of hydrogels spatially and temporally.[56] While UV light is the main method of initiating photopolymerization it has been found that when replaced with visible light, the hydrogel has higher cell compatibility as well as a more uniform structure due to increased penetration depth.[56]

To create a photo ink both Lam *et al.* and Thomas *et al.* combined GelMA and Hyaluronic acid, both of which can be used in extrusion-based bioprinting, and combined with Lithium phenyl-2,4,6-trimethylbenzoyl phosphinate (LAP), a photoinitiator, within phosphate-buffered saline.[60,61] By using a photoinitiator and extrusion-based bioinks, a photoink was synthesized. As shown by these two studies, a photoinitiator is vital to the creation of a photoink. The photoinitiators used most extensively in bioprinting can be classified as free radical photoinitiators. Free radical photoinitiators produce radicals (atom, molecule, or ion with an unpaired valence electron) when exposed to radiation (UV or visible light). Two subcategories of free radical photoinitiators are type 1 and type 2. Currently, the most common option for type 1 photoinitiators is LAP. In contrast, type 2 photoinitiators have a plethora more options such as ruthenium pyridine complex, EY, and camphorquinone.[56]

Disadvantages:

While SLA does have many advantages such as great resolution and low printing time it does have some inherent disadvantages that prevent it from being a perfect bioprinting method. The biggest drawback that SLA faces is the fact that it must use a UV laser to harden the bio-ink, forming a 3D structure. The use of UV light damages the DNA of cells, decreasing their viability. However, this disadvantage is quickly losing its significance as visible light is becoming a more viable option for SLA as demonstrated by Wang *et al.*[57] Another disadvantage that comes with SLA is the extremely long post-processing time. This stage refers to the time when cells are cultured after the initial structure is formed. For example, Grigoryan *et al.* had to culture the hardened cells for up to 15 days.[54] Additionally, Thomas *et al.* cultured their cells for up to 28 days.[60] Thus it can be seen that the high resolution and accuracy seen in SLA can only be achieved after many days of post-processing. Finally, SLA is not the best method for creating multicellular structures.[56] Thus creating complex organs like a heart or liver using SLA is not the most efficient and viable method.

Applications:

While stereolithography does have its disadvantages, the results that can be achieved with SLA more than make up for the disadvantages. For one, SLA has a high resolution of 50 μm[56] with cell viability anywhere between 85% and 90%.[60] In terms of creating vascular structures, SLA has been able to produce viable structures as Thomas *et al.* were able to produce vascular channels using HUVEC and used a degradable photo ink to achieve maximum cell viability of 90% and the endothelial lined channels remained stable for all 28 days they were examined.[60] Stereolithography has no limitations on both bioink viscosity and cell density.[48] Finally, the cost for SLA is quite affordable as Bishop *et al.*[34] described it as "medium". Examples of stereolithography bioprinters include the NGB-R bioprinter and the BIO X printer, priced at $300,000 and $40,000. With all the benefits of stereolithography, there is no doubt that this method holds great potential.

Future Perspectives and Current Limitations:

Currently, bioprinting has unlimited potential and a variety of applications such as organ transplants, artificial meat, and drug testing. However, this subject is still very novel and has some challenges that need to be addressed. One of the challenges that currently exist, according to Zhang *et. al* is the creation of multilayered vascular networks.[63] Vascular structures such as veins and arteries have multiple layers of tissue such as the tunica adventitia and the tunica intima and given that each of these layers has its distinct functions yet also interacts with one another bioprinting must make further advan

DOI: 10.36838/v4i6.17

cements to where a vascular structure can be generated with distinct layers that can also interact with one another. This will further mimic how blood vessels operate within a living organism. One issue with bioprinting that is frequently brought up is the lack of materials within bioink. Researchers are often only limited to synthetic or natural polymers, each with its advantages and disadvantages. Yet, as previously stated, this issue can be solved with the use of hybrid hydrogels which contain the structural integrity of synthetic polymers and the increased cell communication of natural polymers. Finally, arguably the biggest challenge concerning vascular structures specifically in printing resolution. According to Bishop *et. al*, the diameter of a capillary can be as small as 3 μm whereas the highest printing resolution of a bioprinter is 20 μm.[34] What this means is that if a capillary has to be enlarged by nearly 6 times its diameter it means that the rest of the organ must also be enlarged to compensate, resulting in these engineered organs being unusable for human operation. Thus, the priority for bioprinting vascular structures is to increase the printing resolution, and one optimal way to do this could be by combining printing methods, using the high resolution of stereolithography with the speed of extrusion-based bioprinting could greatly increase efficiency while still delivering a viable product.

Table 1: A summary of the differences, advantages, and disadvantages of each of the three printing methods is shown.

Bioprinting Method	Extrusion Based	Droplet-Based	Stereolithography
Bio inks	Hydrogels, Spheroids, Microcarriers, Decellurized ECM	Hydrogels are composed of alginate, fibrin, collagen, and gelatin.	Hydrogels combined with photoinitiators. Ex. GelMA combined with Lithium phenyl-2,4,6-trimethylbenzoyl phosphinate.
Cell Viability	40- 86%[11]	70-90%[39]	85-90%[61]
Disadvantages	The use of a nozzle increases shear stress on cells, decreasing cell viability. This can be mitigated by using hydrogels with a lower viscosity as bioink.	The variety in bio-inks is limited, restricted only to low viscosity bioinks.	The use of ultraviolet light damages cell structure and the post-processing time is quite long, decreasing efficiency.
Maximum Cost of Printer (excluding cost of materials)	$150,000	$200,000	$300,000
Process of Printing	The bio-ink is continuously extruded through a nozzle in layers.	A nozzle generates droplets of bioink that are then placed onto a surface and the droplets then fuse forming a layer.	A light source is used to harden certain areas of a vat of photocurable bioink. After a layer is finished a recoater blade applies a fresh layer of ink to be hardened.
Cells Used	Human umbilical vein endothelial cells [15], human umbilical vein smooth muscle [36], human mesenchymal stem cells [37]	Human mesenchymal stem cells [40], human induced pluripotent stem cells [39], fibroblasts [51,64]	Fibroblasts [58], human umbilical cord endothelial cells [61], lung adenocarcinoma cells [55]
Dimensions	Outer Diameter: 1449 ± 27 μm [36]	Outer Diameter: 4-5 mm [39]	Outer Diameter: 360-720 μm [61]

Conclusion

Bioprinting vasculature remains one of the major challenges that researchers face today when creating synthetic organs. While the fabrication of blood vessels has been successful it has been seen that these synthesized vessels are unable to remain viable for any period longer than one month. This limits their use *in vivo* long-term as they cannot be relied upon to remain functional. Additionally, both veins and arteries are not composed of just one layer of cells. Therefore, another roadblock to the creation of implementable vasculature is the development of multicellular bioprinting, where multiple cell cultures can be placed in a bioink and grow both cohesively and independently, forming layers. This process is already in the works as both multi-material stereolithography and extrusion-based bioprinting have been developed and tested. To create these vasculatures, three bioprinting methods were discussed: Extrusion-based, droplet-based, and stereolithography. As shown in Table 1, while each method does have its drawbacks, all three methods have been able to generate vasculature. In terms of choosing a method for a study, there are multiple factors to consider such as resolution, bio-inks, and cell viability. For resolution and viability, stereolithography is by far the best method and has an affordable cost. Its lack of bio-ink viscosity and cell density only serves to increase its benefits. When on a low budget, droplet-based bioprinting will serve as the best method as out of the three methods it has the lowest cost and has high cell viability at the cost of limited bio-inks. While extrusion-based bioprinting is the most used method and its bioink variety does make it a serviceable method, the lower cell viability makes it much more of a risk to take on. Thus, through our findings, we have found that stereolithography shows the most potential in the creation of functional vasculature.

Acknowledgments

I would like to thank my mentor Nikhith Kalkunte, a Ph.D. student at the University of Texas at Austin for introducing me to the process of academic research and teaching me about the core concepts of regenerative medicine.

References

1. Ozbolat, I. T. Scaffold-Based or Scaffold-Free Bioprinting: Competing or Complementing Approaches? J. *Nanotechnol. Eng. Med.* **2015**, 6 (2), 024701. https://doi.org/10.1115/1.4030414.
2. lghuwainem, A.; Alshareeda, A. T.; Alsowayan, B. Scaffold-Free 3-D Cell Sheet Technique Bridges the Gap between 2-D Cell Culture and Animal Models. *Int. J. Mol. Sci.* **2019**, 20 (19), 4926. https://doi.org/10.3390/ijms20194926.
3. Lee, J. M.; Ng, W. L.; Yeong, W. Y. Resolution and Shape in Bioprinting: Strategizing towards Complex Tissue and Organ Printing. *Appl. Phys. Rev.* **2019**, 6 (1), 011307. https://doi.org/10.1063/1.5053909.
4. lberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P. Blood Vessels, and Endothelial Cells. *Mol. Biol. Cell 4th Ed.* **2002**.
5. Classification & Structure of Blood Vessels | SEER Training https://training.seer.cancer.gov/anatomy/cardiovascular/blood/ classification.html (accessed 2021 -08 -29).
6. Wang, Y.; Zhao, S. Vasculogenesis and Angiogenesis of Human Placenta; Morgan & Claypool Life Sciences, 2010.
7. Ucuzian, A. A.; Gassman, A. A.; East, A. T.; Greisler, H. P. Molecular Mediators of Angiogenesis. J. Burn Care Res. Off. *Publ. Am. Burn Assoc.* **2010**, 31 (1), 158. https://doi.org/10.1097/BCR.0b013e3181c7ed82.
8. Santos-Oliveira, P.; Correia, A.; Rodrigues, T.; Ribeiro-Rodrigues, T. M.; Matafome, P.; Rodríguez-Manzaneque, J. C.; Seiça, R.; Girão, H.; Travasso, R. D. M. The Force at the Tip - Modelling Tension and Proliferation in Sprouting Angiogenesis. *PLOS Comput. Biol.* **2015**, 11 (8), e1004436. https://doi.org/10.1371/journal.pcbi.1004436.
9. Spiegelaere, W. D.; Casteleyn, C.; Broeck, W. V. den; Plendl, J.; Bahramsoltani, M.; Simoens, P.; Djonov, V.; Cornillie, P. Intus susceptive Angiogenesis: A Biologically Relevant Form of An

giogenesis. *J. Vasc. Res*. **2012**, 49 (5), 390–404. https://doi.org/10.1159/000338278.
10. Sweeney, M. D.; Ayyadurai, S.; Zlokovic, B. V. Pericytes of the Neurovascular Unit: Key Functions and Signaling Pathways. *Nat. Neurosci*. **2016**, 19 (6), 771–783. https://doi.org/10.1038nn.4288.
11. Tomasina, C.; Bodet, T.; Mota, C.; Moroni, L.; Camarero-Espinosa, S. Bioprinting Vasculature: Materials, Cells, and Emergent Techniques. *Materials* **2019**, 12 (17), 2701. https://doi.org/10.3390/ma12172701.
12. Ramesh, S.; Harrysson, O. L. A.; Rao, P. K.; Tamayol, A.; Cormier, D. R.; Zhang, Y.; Rivero, I. V. Extrusion Bioprinting: Recent Progress, Challenges, and Future Opportunities. *Bioprinting* **2021**, 21, e00116. https://doi.org/10.1016/j.bprint.2020.e00116.
13. Ning, L.; Yang, B.; Mohabatpour, F.; Betancourt, N.; Sarker, M. D.; Papagerakis, P.; Chen, X. Process-Induced Cell Damage: Pneumatic versus Screw-Driven Bioprinting. *Biofabrication* **2020**, 12 (2), 025011. https://doi.org/10.1088/1758-5090/ab5f53.
14. Willson, K.; Ke, D.; Kengla, C.; Atala, A.; Murphy, S. V. Extrusion-Based Bioprinting: Current Standards and Relevancy for Human-Sized Tissue Fabrication. In 3D Bioprinting: Principles and Protocols; Crook, J. M., Ed.; Methods in Molecular Biology; Springer US: New York, NY, 2020; pp 65–92. https://doi.org/10.1007/978-1-0716-0520-2_5.
15. Suntornnond, R.; Tan, E. Y. S.; An, J.; Chua, C. K. A Highly Printable and Biocompatible Hydrogel Composite for Direct Printing of Soft and Perfusable Vasculature-like Structures. *Sci. Rep*. **2017**, 7 (1), 16902. https://doi.org/10.1038/s41598-017-17198-0.
16. Noh, I.; Kim, N.; Tran, H. N.; Lee, J.; Lee, C. 3D Printable Hyaluronic Acid-Based Hydrogel for Its Potential Application as a Bio ink in Tissue Engineering. *Biomater. Res*. **2019**, 23 (1), 3. https://doi.org/10.1186/s40824-018-0152-8.
17. Lee, B. H.; Lum, N.; Seow, L. Y.; Lim, P. Q.; Tan, L. P. Synthesis and Characterization of Types A and B Gelatin Methacryloyl for Bioink Applications. *Materials* **2016**, 9 (10), 797. https://doi.org/10.3390/ma9100797.
18. Tabriz, A. G.; Hermida, M. A.; Leslie, N. R.; Shu, W. Three-Dimensional Bioprinting of Complex Cell Laden Alginate Hydrogel Structures. *Biofabrication* **2015**, 7 (4), 045012. https://doi.org/10.1088/1758-5090/7/4/045012.
19. Mironov, V.; Visconti, R. P.; Kasyanov, V.; Forgacs, G.; Drake, C. J.; Markwald, R. R. Organ Printing: Tissue Spheroids as Building Blocks. *Biomaterials* **2009**, 30 (12), 2164–2174. https://doi. org/10.1016/j.biomaterials.2008.12.084.
20. Aguilar, I. N.; Olivos, D. J.; Brinker, A.; Alvarez, M. B.; Smith, L. J.; Chu, T.-M. G.; Kacena, M. A.; Wagner, D. R. Scaffold-Free Bioprinting of Mesenchymal Stem Cells Using the Regenova Printer: Spheroid Characterization and Osteogenic Differentiation. *Bioprinting* **2019**, 15, e00050. https://doi.org/10.1016/j.bprint.2019.e00050.
21. Pati, F.; Jang, J.; Ha, D.-H.; Won Kim, S.; Rhie, J.-W.; Shim, J.-H.; Kim, D.-H.; Cho, D.-W. Printing Three-Dimensional Tissue Analogues with Decellularized Extracellular Matrix Bioink. Nat. *Commun*. **2014**, 5 (1), 3935. https://doi.org/10.1038/ncomms4935.
22. Catoira, M. C.; Fusaro, L.; Di Francesco, D.; Ramella, M.; Boccafoschi, F. Overview of Natural Hydrogels for Regenerative Medicine Applications. *J. Mater. Sci. Mater. Med*. **2019**, 30 (10), 115. https://doi.org/10.1007/s10856-019-6318-7.
23. Yahia, Lh. History and Applications of Hydrogels. *J. Biomed. Sci. 2015*, 04 (02). https://doi.org/10.4172/2254-609X.100013.
24. Mancha Sánchez, E.; Gómez-Blanco, J. C.; López Nieto, E.; Casado, J. G.; Macías-García, A.; Díaz Díez, M. A.; Carrasco-Amador, J. P.; Torrejón Martín, D.; Sánchez-Margallo, F. M.; Pagador, J. B. Hydrogels for Bioprinting: A Systematic Review of Hydrogels Synthesis, Bioprinting Parameters, and Bioprinted Structures Behavior. *Front. Bioeng. Biotechnol*. **2020**, 8, 776. https://doi.org/10.3389/fbioe.2020.00776.
25. Madduma-Bandarage, U. S. K.; Madihally, S. V. Synthetic Hydrogels: Synthesis, Novel Trends, and Applications. *J. Appl. Polym. Sci.* **2021**, 138 (19), 50376. https://doi.org/10.1002/app.50376.
26. Palmese, L. L.; Thapa, R. K.; Sullivan, M. O.; Kiick, K. L. Hybrid Hydrogels for Biomedical Applications. Curr. Opin. *Chem. Eng.* **2019**, 24, 143–157. https://doi.org/10.1016/j.coche.2019.02.010.
27. Gyles, D. A.; Castro, L. D.; Silva, J. O. C.; Ribeiro-Costa, R. M. A. Review of the Designs and Prominent Biomedical Advances of Natural and Synthetic Hydrogel Formulations. *Eur. Polym. J*. **2017**, 88, 373–392. https://doi.org/10.1016/j.eurpolymj.2017.01.027.
28. Diniz, I. M. A.; Chen, C.; Xu, X.; Ansari, S.; Zadeh, H. H.; Marques, M. M.; Shi, S.; Moshaverinia, A. Pluronic F-127 Hydrogel as a Promising Scaffold for Encapsulation of Dental-Derived Mesenchymal Stem Cells. *J. Mater. Sci. Mater. Med. 2015*, 26 (3), 153.https://doi.org/10.1007/s10856-015-5493-4.
29. Ozbolat, I. T.; Hospodiuk, M. Current Advances and Future Perspectives in Extrusion-Based Bioprinting. *Biomaterials* **2016**, 76, 321–343. https://doi.org/10.1016/j.biomaterials.2015.10.076.
30. Mehesz, A. N.; Brown, J.; Hajdu, Z.; Beaver, W.; da Silva, J.; Visconti, R.; Markwald, R.; Mironov, V. Scalable Robotic Biofabrication of Tissue Spheroids. *Biofabrication* **2011**, 3 (2), 025002. https://doi.org/10.1088/1758-5082/3/2/025002.
31. Kosheleva, N. V.; Efremov, Y. M.; Shavkuta, B. S.; Zurina, I. M.; Zhang, D.; Zhang, Y.; Minaev, N. V.; Gorkun, A. A.; Wei, S.; Shpichka, A. I.; Saburina, I. N.; Timashev, P. S. Cell Spheroid Fusion: Beyond Liquid Drops Model. *Sci. Rep*. **2020**, 10 (1), 12614. https://doi.org/10.1038/s41598-020-69540-8.
32. Yang, J.; Guertin, P.; Jia, G.; Lv, Z.; Yang, H.; Ju, D. Large-Scale Microcarrier Culture of HEK293T Cells and Vero Cells in Single-Use Bioreactors. *AMB Express* **2019**, 9, 70. https://doi. org/10.1186/s13568-019-0794-5.
33. Levato, R.; Visser, J.; Planell, J. A.; Engel, E.; Malda, J.; Mateos-Timoneda, M. A. Biofabrication of Tissue Constructs by 3D Bioprinting of Cell-Laden Microcarriers. *Biofabrication* **2014**, 6 (3), 035020. https://doi.org/10.1088/1758-5082/6/3/035020.
34. Bishop, E. S.; Mostafa, S.; Pakvasa, M.; Luu, H. H.; Lee, M. J.; Wolf, J. M.; Ameer, G. A.; He, T.-C.; Reid, R. R. 3-D Bioprinting Technologies in Tissue Engineering and Regenerative Medicine: Current and Future Trends. *Genes Dis.* **2017**, 4 (4), 185–195. https://doi.org/10.1016/j.gendis.2017.10.002.
35. Hinton, T. J.; Jallerat, Q.; Palchesko, R. N.; Park, J. H.; Grodzicki, M. S.; Shue, H.-J.; Ramadan, M. H.; Hudson, A. R.; Feinberg, A. W. Three-Dimensional Printing of Complex Biological Structures by Freeform Reversible Embedding of Suspended Hydrogels. *Sci. Adv.* **2015**, 1 (9), e1500758. https://doi.org/10.1126/sciadv.1500758.
36. Zhang, Y.; Yu, Y.; Akkouch, A.; Dababneh, A.; Dolati, F.; T. Ozbolat, I. In Vitro Study of Directly Bioprinted Perfusable Vascular Conduits. *Biomater. Sci.* **2015**, 3 (1), 134–143. https://doi.org/10.1039/C4BM00234B.
37. Jia, W.; Gungor-Ozkerim, P. S.; Zhang, Y. S.; Yue, K.; Zhu, K.; Liu, W.; Pi, Q.; Byambaa, B.; Dokmeci, M. R.; Shin, S. R.;Khademhosseini, A. Direct 3D Bioprinting of Perfusable Vascular Constructs Using a Blend Bioink. *Biomaterials* **2016**, 106, 58–68. https://doi.org/10.1016/j.biomaterials.2016.07.038.
38. 3D bioprinting 2021: guide and product selection https://www.aniwaa.com/buyers-guide/3d-printers/3d-bioprinting-3d-bioprinters/ (accessed 2021 -08 -29).
39. Gudapati, H.; Dey, M.; Ozbolat, I. A Comprehensive Review on Droplet-Based Bioprinting: Past, Present and Future. Biomaterials 2016, 102, 20–42. https://doi.org/10.1016/*j.biomaterials.***2016**.06.

DOI: 10.36838/v4i6.17

012.
40.Jentsch, S.; Nasehi, R.; Kuckelkorn, C.; Gundert, B.; Aveic, S.; Fischer, H. Multiscale 3D Bioprinting by Nozzle-Free Acoustic Droplet Ejection. *Small Methods* **2021**, 5 (6), 2000971. https://doi.org/10.1002/smtd.202000971.
41.Guo, Q.; Su, X.; Zhang, X.; Shao, M.; Yu, H.; Li, D. A Review on Acoustic Droplet Ejection Technology and System. *Soft Matter* **2021**, 17 (11), 3010–3021. https://doi.org/10.1039/D0SM02193H.
42.Ng, W. L.; Lee, J. M.; Yeong, W. Y.; Naing, M. W. Microvalve-Based Bioprinting – Process, Bio-Inks and Applications. Biomater. *Sci*. **2017**, 5 (4), 632–647. https://doi.org/10.1039/C6BM00861E.
43.Ng, W. L.; Yeong, W. Y. MICROVALVE BIOPRINTING OF CELLULAR DROPLETS WITH HIGH RESOLUTION AND CONSISTENCY. 7.
44.Włodarczyk-Biegun, M. K.; del Campo, A. 3D Bioprinting of Structural Proteins. *Biomaterials* **2017**, 134, 180–201. https://doi.org/10.1016/j.biomaterials.2017.04.019.
45.Electrohydrodynamic Printing (E-Jet) – Pikul Research Group.
46.How Piezoelectricity Works | EAGLE | Blog https://www.autodesk.com/products/eagle/blog/piezoelectricity/ (accessed 2021-08-12).
47.Gu, Z.; Fu, J.; Lin, H.; He, Y. Development of 3D Bioprinting: From Printing Methods to Biomedical Applications. Asian J. *Pharm. Sci*. **2020**, 15 (5), 529–557. https://doi.org/10.1016/jajps.2019.11.003.
48.Umezu, S. Precision Printing of Gelatin Utilizing Electrostatic Inkjet. Jpn. J. Appl. Phys. 2014, 53 (5S3), 05HC01. https://doi.org/10.7567/JJAP.53.05HC01.
49.Donderwinkel, I.; Hest, J. C. M. van; Cameron, N. R. Bio-Inks for 3D Bioprinting: Recent Advances and Future Prospects. *Polym. Chem*. **2017**, 8 (31), 4451–4471. https://doi.org/10.1039/C7PY00826K.
50.Yao, R.; Alkhawtani, A. Y. F.; Chen, R.; Luan, J.; Xu, M. Rapid and Efficient in Vivo Angiogenesis Directed by Electro-Assisted Bioprinting of Alginate/Collagen Microspheres with Human Umbilical Vein Endothelial Cell Coating Layer. *Int. J. Bioprinting* **2019**, 5 (2.1), 194. https://doi.org/10.18063/ijb.v5i2.1.194.
51.Xu, C.; Zhang, M.; Huang, Y.; Ogale, A.; Fu, J.; Markwald, R. R. Study of Droplet Formation Process during Drop-on-Demand Inkjetting of Living Cell-Laden Bioink. **2014**, 9.
52.Xu, C.; Huang, Y.; Fu, J.; Markwald, R. R. Electric Field-Assisted Droplet Formation Using Piezoactuation-Based Drop-on- Demand Inkjet Printing. *J Micromech Microeng* **2014**, 10.
53.Cui, X.; Boland, T. Human Microvasculature Fabrication Using Thermal Inkjet Printing Technology. *Biomaterials* **2009**, 30 (31), 6221–6227. https://doi.org/10.1016/j.biomaterials.2009.07.056.
54.Cui, X.; Dean, D.; Ruggeri, Z. M.; Boland, T. Cell Damage Evaluation of Thermal Inkjet Printed Chinese Hamster Ovary Cells. *Biotechnol. Bioeng.* **2010**, 106 (6), 963–969. https://doi.org/10.1002/bit.22762.
55.Grigoryan, B.; Sazer, D. W.; Avila, A.; Albritton, J. L.; Padhye, A.; Ta, A. H.; Greenfield, P. T.; Gibbons, D. L.; Miller, J. S. Development, Characterization, and Applications of Multi-Material Stereolithography Bioprinting. *Sci. Rep*. **2021**, 11 (1), 3171. https://doi.org/10.1038/s41598-021-82102-w.
56.SLA vs DLP: The Differences – Simply Explained https://all3dp.com/2/dlp-vs-sla-3d-printing-technologies-shootout/ (accessed 2021 -08 -05).
57.Zheng, Z.; Eglin, D.; Alini, M.; Richards, G. R.; Qin, L.; Lai, Y. Visible Light-Induced 3D Bioprinting Technologies and Corresponding Bioink Materials for Tissue Engineering: A Review. *Engineering* **2020**. https://doi.org/10.1016/j.eng.2020.05.021.
58.Wang, Z.; Abdulla, R.; Parker, B.; Samanipour, R.; Ghosh, S.; Kim, K. A Simple and High-Resolution Stereolithography-Based 3D Bioprinting System Using Visible Light Crosslinkable Bioinks. *Biofabrication* **2015**, 7 (4), 045009. https://doi.org/10.1088/1758-5090/7/4/045009.
59.All About Stereolithography 3D Printing https://www.thomasnet.com/articles/custom-manufacturing-fabricating/stereolithography-3d-printing (accessed 2021 -09 -12).
60.Stereolithography / SLA 3D Printing – Simply Explained https://all3dp.com/2/stereolithography-3d-printing-simply-explained/ (accessed 2021 -09 -12).
61.Thomas, A.; Orellano, I.; Lam, T.; Noichl, B.; Geiger, M.-A.; Amler, A.-K.; Kreuder, A.-E.; Palmer, C.; Duda, G.; Lauster, R.; Kloke, L. Vascular Bioprinting with Enzymatically Degradable Bioinks via Multi-Material Projection-Based Stereolithography. *Acta Biomater*. **2020**, 117, 121–132. https://doi.org/10.1016/j.actbio.2020.09.033.
62.Lam, T.; Dehne, T.; Krüger, J. P.; Hondke, S.; Endres, M.; Thomas, A.; Lauster, R.; Sittinger, M.; Kloke, L. Photopolymerizable Gelatin and Hyaluronic Acid for Stereolithographic 3D Bioprinting of Tissue-Engineered Cartilage. *J. Biomed. Mater. Res. B Appl. Biomater*. **2019**, 107 (8), 2649–2657. https://doi.org/10.1002/jbm.b.34354.
63.Zhang, Y.; Kumar, P.; Lv, S.; Xiong, D.; Zhao, H.; Cai, Z.; Zhao, X. Recent Advances in 3D Bioprinting of Vascularized Tissues. Mater. Des. 2021, 199, 109398. https://doi.org/10.1016/j.matdes.2020.109398.
64.Ng, W. L.; Yeong, W. Y.; Naing, M. W. Polyvinylpyrrolidone-Based Bio-Ink Improves Cell Viability and Homogeneity during Drop-On-Demand Printing. Materials 2017, 10 (2), 190. https://doi.org/10.3390/ma10020190.

Author

Tarun Rao is a senior at Leland High School in San Jose California. In college he plans to major in Biomedical Engineering. Some activities that he enjoys include playing the violin and playing tennis.

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Effects of Facial Expressions on Willingness to Purchase

Divya Subramanian
The Pingry School, 131 Martinsville Road, Basking Ridge, NJ, 07920, USA; dsubramanian.05@gmail.com

ABSTRACT: In static advertisements, images instead of videos are used to promote products, and it is most often the faces of the models that get the consumers' attention. How facial expressions transfer emotions from the producer to the consumer has been explained by the emotional contagion theory, which suggests the recipient replicates the emitter through motor mimicry which changes the recipient's emotional state. This paper investigates the effects of facial expressions in advertisements on adolescents in comparison to adults. It was hypothesized that participants would give a higher product evaluation if the model displayed a real smile, versus a fake smile or neutral facial expression. It was also hypothesized that adolescents would be more susceptible to facial expressions and therefore pay more for the same product than adults. High school and adult participants completed a survey in which they evaluated a product (a white t-shirt) when the model displayed a neutral face, fake smile, and real smile. The results showed that participants gave higher product evaluation scores and were willing to pay more when the model displayed a genuine smile over the other facial expressions. In addition, adolescents were less influenced by the model's expressions than adults which may be due to their lack of experience in spending. To extend this study, the effects of emotional contagion theory on other factors including how the product attracts attention or lasts in the consumers' memory could be tested.

KEYWORDS: Behavioral and Social Science; Sociology and Social Psychology; Adolescence; Facial Expressions; Purchasing; Emotional Contagion.

Introduction

Every day, American consumers spend $29.3 billion on average.[1] Annually, $296.4 billion US dollars are spent on advertising for these products.[2] Human faces are oftentimes at the center of these marketing campaigns. Adolescents are one of the key targets of advertising perhaps due to their susceptibility to influence.[3] This paper investigates the effects of facial expressions in advertisements on adolescents. For the purpose of this paper, the following definitions will be adopted. First, facial expressions display a person's discrete emotional state. Second, the effect of advertisements can be assessed through the willingness to purchase. The willingness to purchase, as defined in behavioral economics, is the highest price a consumer is willing to pay for a product. Third, adolescents are defined biologically as individuals in the 10 - 19 years age group in a transitional phase of growth and maturation through puberty to adulthood. Adolescence marks a time of substantial change in physical maturation, cognitive abilities, and social interactions.

In static advertisements, images, instead of videos, are used to promote the product. A variety of facial expressions and postures are emphasized by models in order to convey the appropriate message. It is most often the faces of the models that get the consumers' attention.[4] How facial expressions transfer emotions from the producer to the consumer has been explained by the emotional contagion theory.[5] This theory has two steps: at first, the recipient replicates the emitter through motor mimicry. Second, mimicry changes the recipient's emotional state. The emotional contagion theory has, in marketing, been used in face-to-face interaction. This research has supported the idea that product evaluation increases with a genuine smile.[6] How a smile's intensity affects perception of competence in static ads have also been studied but this did not address product evaluation.[7] While different types of smiles have been studied, this paper investigates the effects of a fake smile compared to a genuine smile and a neutral facial expression on product evaluation.

In this study, a smile is defined as when a person exemplifies a positive emotion in which the corners of their mouths are typically turned up. A Duchenne, or real, smile involves displaying a smile which accurately displays what the person feels. This often includes scrunching the eyes, known as smizing. On the other hand, a fake smile only uses the zygomatic major muscle. For example, a Duchenne smile is the smile of a championship winner while the runner up typically displays a fake smile. [8] Based on these variances, a consumer should notice the difference between genuine and fake smiles.

This paper expands on recent findings by Isabella and Vieira who investigated emotional contagion effects in advertisements.[8] This denotes the process in which emotions and related behavior are spread from sender to receiver. One of their experiments tested the influence of neutral, fake, and genuine smiles of a model on the product evaluation. The experiments did in fact prove that the facial expressions influenced the product evaluation. They showed that genuine smiles tend to increase product evaluation over false smiles because of the mimicry response.

The brain is not fully developed until 25. Pechmann studied how adolescents are more impulsive and self-conscious

DOI: 10.36838/v4i6.18

than adults and are therefore more susceptible to fall for advertisements.[3] Especially with the growing prevalence of social media and technology on youth, it is important to see how the behavior of the new generation is affected, if at all, by facial expressions.

Based on the previous research of facial expressions on advertisements outlined above, the hypotheses for the present study are the following:

Hypothesis 1: All participants will be most susceptible to a genuine smile over a fake smile and a fake smile over a neutral facial expression.

Hypothesis 2: Adolescents will be more susceptible to advertisements and therefore will be more willing to pay a higher price for the same product compared to adults.

Hypothesis 3: Adolescents will be more susceptible to facial expressions and therefore a genuine smile will produce an even more positive influence in product evaluation for adolescents than adults.

Methods

In this study, the participants were separated into two groups: adults [n= 30] and adolescents [n= 20]. The survey was administered to an online panel of adults older than 19 in the fall of 2021. 99 responses were received, and 30 were randomly selected for the analysis so that the sample of adolescents and adults had balanced numbers. The same survey was distributed at a high school in New Jersey.

The procedure of the present study replicated that of Isabella and Viera.[8] A survey was conducted that showed pictures of a female face with a neutral expression, fake smiling, and genuinely smiling (Figure 1) in a static advertisement. In this survey, the product being sold was a plain white t-shirt. Before measuring the willingness to purchase, a control measure was included in the survey to check for congruence between the model's facial expression and the participant's interpretation of it. This control was used to eliminate participants. Twenty-nine (29) adults and 3 adolescents were not included in the results because they did not accurately perceive the model's facial expression. Participants then rated their appeal and judgment of the product based on questions from the Beren's Product evaluation scale, which asks participants to rate the favorability, their likelihood of purchasing the product, and the preferred pricing of the product on a 7-point Likert scale.[9] The XLMiner Analysis ToolPak was then used to conduct ANOVA and t-tests on these results.

Figure 3: Neutral, fake, and genuine facial expressions, respectively

Methods

In total, the survey responses from 30 adult participants were randomly selected and analyzed. Of these, 26.67% were female and 73.33% were male. 73.33% of participants identified as Caucasian, 10% identified as African American, and 23.33% identified as Asian American or Pacific Islander.

In total, 20 adolescent participants answered the survey. Of these, 75% were female, 20% were male, and 5% preferred not to disclose their gender. 50% of participants identified as Caucasian, 5% identified as African American, 45% identified as Asian American or Pacific Islander, and 5% identified as Hispanic. The survey was an online survey sent to a high school and participants were included in an optional raffle as an incentive for participation.

The two main aspects of the survey that were analyzed were the adult and adolescent participants' price evaluation and judgment on the quality of the product. The participants were asked to consider both the product and advertiser to determine what price, in dollars, they would pay for the product. To judge quality, the participants were asked how favorable their judgment was of this product and were asked to consider the quality of the white shirt in comparison to similar products on a scale from 1 to 7.

ANOVA Results:

Single factor ANOVAs were conducted to compare the responses of both adults and adolescents for both quality and preferred pricing across the three faces: neutral, fake smile, real smile. On average, there was a statistically significant difference in adults' rating of product quality across a neutral face (M= 2.53, SD= 1.48), fake smile (M= 4.37, SD= 1.47), and real smile (M= 4.2, SD= 2.02); $F(2, 87)= 10.94$, $p< 0.0001$ (Figure 2.1). In addition, there was a statistically significant difference in adults' ratings for preferred pricing across a neutral face (M = 6.67, SD= 5.55), fake smile (M= 9.6, SD= 6.67), and real smile (M= 11.13, SD= 8.61); $F(2, 87)= 3.1$, $p< 0.05$ (Figure 2.2). On average, there was a statistically significant difference in adolescents' ratings of product quality across the neutral face (M= 2.55, SD= 1.50), fake smile (M= 3.8, SD= 1.48), and real smile (M= 3.9, SD= 2.02) conditions; $F(2, 57) = 3.98$, $p< 0.025$ (Figure 2.3). Lastly, there was a statistically significant difference in adolescents' ratings for preferred pricing across a neutral face (M= 6.92, SD= 2.89), fake smile (M= 9.32, SD= 3.91), and real smile (M= 10.17, SD= 3.96); $F(2, 57)= 4.33$, $p < 0.018$ (Figure 2.4).

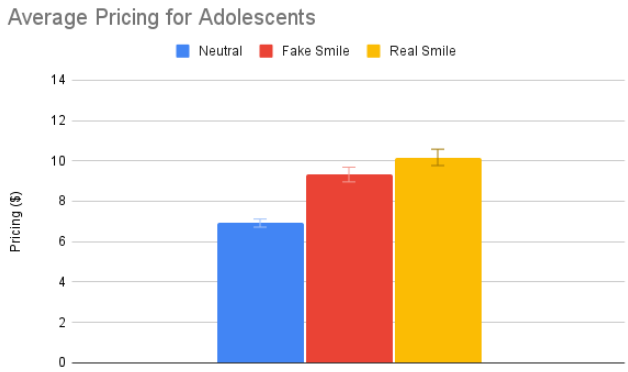


Figure 2.1: Average quality rating on Beren's scale for adults across three facial expressions.

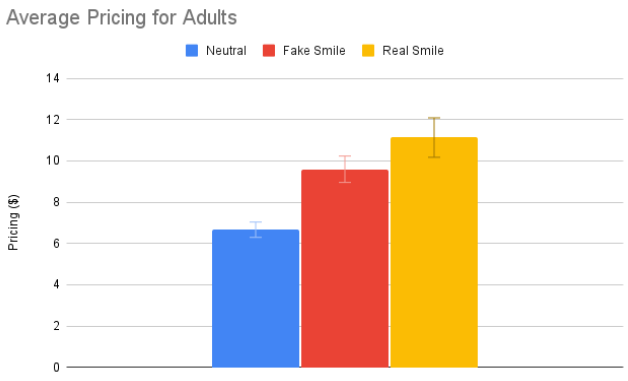


Figure 2.2: Average product pricing for adults across the three facial expressions.

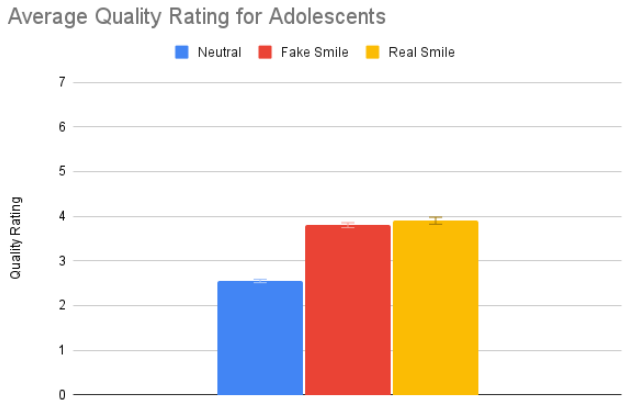


Figure 2.3: Average quality rating on Beren's scale for adolescents across three facial expressions.

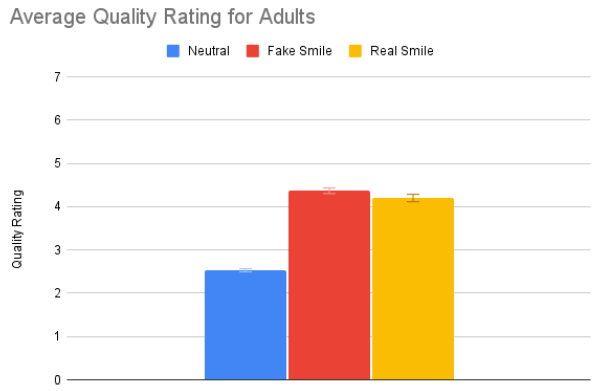


Figure 2.4: Average product pricing for adolescents across three facial expressions.

Paired Sample T-Tests

Adults:

To further explore the results from the ANOVA analysis, paired sample t-tests were used to determine which facial expression conditions were in fact statistically different. According to the results of the Paired Sample T-Test, adults' ratings of quality across a neutral face and fake smile ($p< 0.00001$) as well as across a neutral face and real smile ($p< 0.001$) were both statistically significant. On the other hand, adults' ratings across a fake smile and real smile were not statistically significant ($p= 0.54$).

The adults' ratings of product pricing across a neutral face and fake smile ($p< 0.002$), neutral face and real smile ($p< 0.003$), and fake smile and real smile ($p< 0.03$) were all statistically significant.

Adolescents:

According to the results of the Paired Sample T-Test, adolescents' ratings of quality across a neutral face and fake smile ($p< 0.006$) as well as across a neutral face and real smile ($p< 0.02$) were both statistically significant. On the other hand, adolescents' ratings across a fake smile and real smile were not statistically significant ($p> 0.05$).

Adolescents' ratings of product pricing across a neutral face and fake smile ($p< 0.0002$), neutral face and real smile ($p< 0.0002$), and fake smile and real smile ($p< 0.04$) were all statistically significant.

The findings support previous research on emotional contagion in which happiness exemplified by the advertiser evokes that same emotion in the participants. In all cases, adolescents and adults were on average willing to pay more and had a significantly higher evaluation of the quality of the product when the advertiser was smiling rather than displaying a neutral facial expression.

The results show that participants placed higher evaluation scores of product pricing when presented with an advertiser with a genuine smile, supporting hypothesis 1. To explain these results, previous studies can be examined, which have revealed that the muscular contractions involved in smiling and fake smiling can actually underlie how a viewer regards the expression and how positive emotions are transferred.[10,11] With a genuine smile, the *zygomatic major* and *orbicularis oculi* muscles in the face are contracted. As a consequence, people tend to intuitively evaluate a smiling facial expression as authentic, genuine, and trustworthy.[12] In the present study, it is possible that the advertisers' genuine smile inspired a sense of trust and positivity in the participants viewing the product, leading to higher product quality ratings in the smiling condition. On the other hand, a fake smile only involves the zygomatic major muscle.[13] The data shows that consumers can subconsciously observe the difference between a genuine and false smile and be affected differently. In the present study, it appears that participants responded differently to a genuine versus fake smile as their product quality ratings and price ratings differed.

It was hypothesized that adolescents would pay a greater amount than adults for the same product due to their lack of experience earning and spending money. The results of this study surprisingly showed that adolescents overall had a lower evaluation of the product, negating hypothesis 2. For both product quality and preferred pricing for all three facial expressions, adolescents, on average, gave a lower rating or pricing for each category (see figure 2 to compare these averages) compared to adults. For adolescents, the greater the susceptibility to interpersonal influence, the greater the tendency to buy on impulse.[14] Therefore, when shown a model with a genuine smile, adolescents were more willing to pay a higher price and give the product a better evaluation, which is also demonstrated by the results. However, adolescents were significantly less influenced by a false versus a real smile, opposing hypothesis

DOI: 10.36838/v4i6.18

3, when pricing the product compared to adults. On average, adults were willing to pay about 14.6% more when the model had a real smile while adolescents were willing to pay about 9.12% more. This can be attributed to the lack of exposure of adolescents to spending, which results in an unformed understanding of the value of money. This could be explained since adolescents will typically have less money leading to cheaper spending tendencies. Part of this study involved having participants imagine that this was a real product being sold. In other words, they had to envision that this still face on the screen was associated with selling this white t-shirt. Perhaps adults were better able to envision this scenario, while adolescents struggled to make this link.

Conclusion

In conclusion, consumers are willing to pay more and tend to have a higher evaluation of the quality of a product when a model in an advertisement displays a real smile, instead of a fake smile or a neutral facial expression in marketing material. Adults have a better evaluation and are willing to pay more than adolescents for the same product. Lastly, adults are more susceptible to facial expressions since their price evaluation when presented with a fake smile to a real smile changed more dramatically than that of adolescents. These implications suggest marketing managers would benefit from advertising with genuine smiles since they will evoke positive emotions in the consumer via emotional contagion, making them more likely to purchase.

The study could have benefitted from a larger sample size to increase the amount of data collected. This would improve the validity and generalizability of the data. Future investigations could explore the effects of emotional contagion on other factors including how the product attracts attention or lasts in the consumers' memory. A more detailed survey could allow researchers to determine what specific emotions are elicited from different facial expressions. Different stimuli such as the race or gender of the advertiser instead of facial expression could be used to measure product evaluation. In addition, this study had varied gender distributions between adolescents and adults which could have impacted the results. This study could be replicated with a more even sampling in order to test within group differences. This study researched the effects of facial expressions in static advertisements, and future research could also replicate this with dynamic advertisements (i.e., moving images for instance).

Acknowledgments

Thank you Florentine Salmony and Hannah Dunn for mentoring and helping me with my project.

References

1. Brittany De Lea. (2019, March 12). How much the average American spends per day. Fox Business; Fox Business https://www.foxbusiness.com/personal-finance/how-much-the-average-american-spends-per-day
2. North America ad spend 2024 | Statista. (2021). Statista; Statista. https://www.statista.com/statistics/429036/advertising-expenditure-in-north-america/
3. Pechmann C, Levine L, Loughlin S, Leslie F. Impulsive and Self-Conscious: Adolescents' Vulnerability to Advertising and Promotion. Journal of Public Policy & Marketing. 2005;24(2):202-221. https://doi.org/10.1509/jppm.2005.24.2.202
4. Bindemann, M., Burton, A. M., Langton, S. R. H., Schweinberger, S. R., & Doherty, M. J. (2007). The control of attention to faces. Journal of Vision, 715, 1–8. https://doi.org/10.1167/7.10.15
5. Dallimore, K. S., Sparks, B. A., & Butcher, K. (2007). The influence of angry customer outbursts on service providers' facial displays and affective states. Journal of Service Research, 10, 78–92. https://doi.org/10.1177/1094670507304694
6. Howard, D. J., & Gengler, C. (2001). Emotional contagion effects on product attitudes. Journal of Consumer Research, 28, 189–201. https://doi.org/10.1086/322897
7. Wang, Q., Xu, Z., Cui, X., Wang, L., & Ouyang, C. (2016). Does a big duchenne smile really matter on e-commerce websites? An eye-tracking study in China. Electronic Commerce Research, 17,1-18. https://doi.org/10.1007/s10660-016-9237-4
8. Isabella, G., & Vieira, V.A. (2020). The effect of facial expression on emotional contagion and product evaluation in print advertising. RAUSP Management Journal, 55, 375-391. https://doi.org/10.1108/RAUSP-03-2019-0038
9. Berens, G., Van Riel, C. B., & Van Bruggen, G. H. (2005). Corporate associations and consumer product responses: The moderating role of corporate brand dominance. Journal of Marketing, 69, 35–48.https://doi.org/10.1509/jmkg.69.3.35.66357
10. Hennig-Thurau, T., Groth, M., Paul, M., & Gremler, D. D. (2006). Are all smiles created equal? How emotional contagion and emotional labor affect service relationships. Journal of Marketing, 70, 58-73. https://doi.org/10.1509/jmkg.70.3.58
11. Howard, D. J., & Gengler, C. (2001). Emotional contagion effects on product attitudes. Journal of Consumer Research, 28, 189–201. https://doi.org/10.1086/322897
12. Gunnery, S. D., & Ruben, M. A. (2015). Perceptions of duchenne and non-duchenne smiles: A metaanalysis. Cognition and Emotion, 9931, 1–15. https://doi.org/10.1080/02699931.2015.1018817
13. Messinger, D. S., Mattson, W. I., Mahoor, M. H., & Cohn, J. F. (2012). The eyes have it: Making positive expressions more positive and negative expressions more negative. Emotion, 12, 430–436. https://doi.org/10.1037/a0026498
14. Lin, Yi-Hsiu, and Chen-Yueh Chen. "Adolescents' impulse buying: susceptibility to interpersonal influence and fear of negative evaluation." Social Behavior and Personality: An International Journal, vol. 40, no. 3, Apr. 2012, pp. 353+. https://doi.org/10.2224/sbp.2012.40.3.353

Author

Divya Subramanian is a junior at The Pingry School who thoroughly enjoys the fields of math, science, and computer science. She is in her school's research team which uses computer science to code evolutionary models of neural networks. In the future, she hopes to combine her interests in math and psychology to pursue a career in marketing and behavioral economics.

IJHSR
ijhighschoolresearch.org
REVIEW ARTICLE

Understanding the Role of Physical Factors on the Epigenetic Control of Type 2 Diabetes in Tamil Nadu, India

Jerosh A. Susaiprakash
Emirates International School Jumeirah, Dubai, United Arab Emirates; jeroshantoine@gmail.com

ABSTRACT: Type 2 Diabetes is an exponentially growing condition affecting the lives of entire populations. The significant prevalence of cases raises issues in the healthcare industry regarding funding and the efficacy of treatments. The magnitude of the problem calls for new insights into epigenetics and whether strategic changes in physical activity can alter the epigenome to become favorable in the prevention and treatment of cases. Epigenetic changes produced by acute exercise or resistance training groups link directly to outcomes connected with the incidence of Type 2 diabetes. Individuals who engage in rigorous and high-intensity exercise lower their chance of acquiring Type 2 diabetes and reduce the condition's symptoms and effects.

KEYWORDS: Biomedical and Health Sciences; Genetics and Molecular Biology of Disease; Epigenetics; Type 2 Diabetes; Physique and Obesity.

Introduction

Diabetes and its significant prevalence have always been a common topic of discussion while growing up in Tamil Nadu, India. From 2014 to 2017, the prevalence of Diabetes Mellitus rose from 6% to 8.3%.[1] This was quite interesting as a student studying abroad as Type 2 Diabetes was prevalent in countries outside India, however, there was a noticeable difference in awareness within South-Indian communities. This paper was written to understand this observation and any links to epigenetics, a relatively new scientific study governing genomes' adaptations and lifestyle changes. This study could aid in developing new concepts on lifestyle changes that can be introduced into the South-Indian communities examined in this study to reduce their risk of Type 2 Diabetes. The study focuses on Type 2 diabetes as it links to dietary and lifestyle causes.

Discussion

Prerequisite Knowledge:

Type 2 Diabetes is a chronic condition where an individual develops insulin resistance. Insulin is an essential hormone produced by the body to signal the storage of glucose. Glucose is the monomer of carbohydrates that are consumed as food and transported around the circulatory system. When blood glucose levels are high, the pancreas produces insulin to store the excess glucose; however, an individual experiencing Type 2 diabetes would be inefficient at this process. As a result, their cells become desensitized to the hormone, and an excess of glucose remains in the circulatory system (hyperglycemia). Thus, the issue is not the lack of insulin; it is the insulin's inefficacy to signal cells to increase glucose uptake from the blood.[2]

Traditionally, Type 1 diabetes links to genetic predisposition and occurrences within family history. Although having no specific causes, identifying such variants of the HLA-DQB, HLA-DRB1, and HLA-DQA1 genes were observed to increase the likelihood of developing the condition known as diabetes. These genes code for the pancreatic beta cells to develop inappropriate immunological responses. The result of the responses leads to the destruction and impairment of pancreatic beta cells. Thus, there is an insulin secretion impairment and can be considered an autoimmune disorder as well.[3]

The genome determines our bodies' physiological and bio-chemical structure; all the genetic information is present in chromosomes. However, looking closer into the structure of a single strand of Deoxyribonucleic Acid (DNA), it can be observed that the double-helical structure consists of repeating units of nucleotides. Nucleotides are composed of a phosphate group, deoxyribose (pentose sugar), and a nitrogenous base. The complementary base pairing between pyrimidines (Thymine and Cytosine) and purines (Adenine and Guanine) leads to coding specific genes. These genetic combinations can remain for life however can be permanently altered in the case of genetic damage caused by carcinogens and ionizing radiation. A relatively young research field on the epigenome has showcased the importance of environmental and lifestyle factors in altering the ability of the body to read the genetic code. Characterized by the ability of specific genes to be expressed, epigenetic factors affect the role of genes in the pathologies of conditions as well.[4]

Specific processes occur on the molecular level to activate or repress specific genes. Several modifications such as methylation and acetylation occur to cause these changes. For example, the addition of a methyl group to a nucleosome represses the genetic expression by interfering with the ability of transcription to bind with a strand of DNA.[5] External stimuli in the environment induce changes in transcription factor levels using chemical responses. As a result, the degree of gene expression varies. A notable example would be Hydrangeas (a flowering plant) that blooms with blue flowers in acidic soils and pink flowers in alkaline soils. In humans, ex-

DOI: 10.36838/v4i6.19

posure to solar radiation influences the production of melanin in the epidermal layer.[6]

Issue in Study:

Tamil Nadu is a large state in India with a population of over 72.15 million people.[7] In a 2010 study, it was found that roughly 6% of the studied group were diabetic, and 17.2% of the subset had a Body Mass Index (BMI) of greater than 30 (obese).[8] This observation was commonly noted within the South-Indian population and hence, the researchers of the study proposed Indian Diabetic Risk Score (IDRS) parameter to better understand the subset population's vulnerability to Type 2 Diabetes. The IDRS is a score out of 100 which takes into account age, abdominal obesity, family history of diabetes, and physical activity. Active diabetic history and increasing age and abdominal obesity are linked to a higher IDRS whilst increasing physical activity lowered the IDRS. The study followed up 1936 individuals from February to April, 2018.[8] Although BMI was adopted as a universal measure of physique, ethnic group-based BMI charts were introduced and adopted. BMI is a calculated score that is used to determine body fat levels using the following equation:

$$Body\ Mass\ Index = \frac{Weight\ in\ Kilograms}{Height\ Squared\ \text{in}\ Meters}$$

Example Calculation:

The average Indian man weighs 65 kg and has a height of 177 cm.[9]

$$Body\ Mass\ Index = \frac{65}{1.77^2} = 20.7$$

According to the World Health Organization's Asian BMI categories, the average man would be classified as Ideal Weight (18.5 – 23).[10]

It is commonly known that high obesity rates increase the risk of conditions such as Type 2 diabetes and atherosclerosis. However, less than a fifth of the diabetic set was classified as obese. Although BMI has had a reputation for neglecting the role of lean body mass on the score, the data suggests that the more significant percentage (82.8%) were significantly leaner.[7] This contributes to the notion that BMI alone is not a driving factor in solving the Type 2 diabetes epidemic in Tamil Nadu.[11]

Role of Occupational Activity Levels:

An understanding of the socio-demographic distribution of Tamil Nadu sourced from an external study can be attained using Table 1. This study covered 35,000 individuals from rural health centers in Chunampett and Annechikuppam from February to March, 2008.[7]

Table 1: Occupational Classification of Studied Population

Occupation	Number	Percentage of Total Study (1196)
Homemaker	931	48.14
Labor work	292	15.08
Business	118	6.10
Agriculture	98	5.06
Retired	51	2.63
Student	44	2.27
Service	35	1.81
Other	366	19

Working with the assumption that obesity levels are not the primary cause of Type 2 Diabetes, it is possible to delve into the activity levels of these occupations using Table 2.[12]

Table 2: Occupational Activity Level (OA)

Occupation	Activity Level
Homemaker	Intermediate (110 OA)
Labor work	High (36 OA)
Business	Intermediate (156 OA)
Agriculture	High (31 OA)
Retired	-
Student	Low (183 OA)
Service	High (44 OA)
Other	-

Note: Lower score indicates a more active occupational activity level based on accelerometer data.

Excluding the two null groups (Retired and Other), 33 individuals fall under the 'Low OA' category, 1049 individuals fall under the 'Intermediate OA' category, and 425 individuals fall under the 'High OA' category.

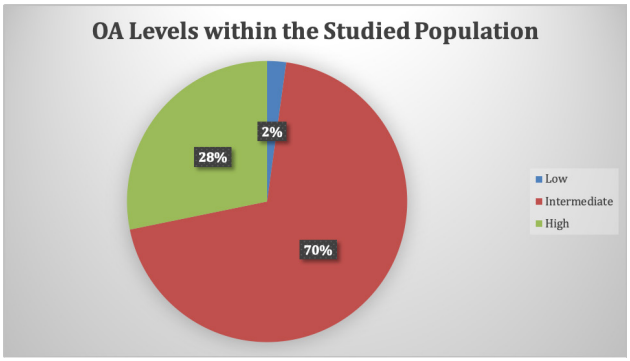


Figure 1: Pie Chart displaying OA Levels within the Studied Population.[12]

Using Figure 1, the vast percentage (70%) of the demographic lies within the Intermediate OA level. The data from Table 2 were gathered using accelerometry, a technique used to measure the mechanical movement of the subject. On paper, this should be pretty accurate for understanding the activity levels. However, there needs to be a clear division between metabolic activity and muscular hypertrophy. Although activities such as walking accumulate to higher metabolic activity levels, they lack sufficient myofibrillar strain, which fails to achieve hypertrophy. Hypertrophy is crucial in developing more leaner mass which acts as a 'glucose sink'. Myofibrils are a basic unit of striated muscle cells responsible for the tensile strength of the cell. Under mechanical tension from activities such as resistance training, these fibers tend to get damaged. Responsive measures from the body repair and build more fibers to resist the strain better. Muscle cells are also multi-nucleated as they are required to generate power for mechanical motion, e.g., flexing an arm. This process requires metabolic respiration to convert glucose into Adenosine Tri-Phosphate (ATP), of which the glucose is absorbed through the bloodstream. This explains the significance of higher lean body mass to reduce hyperglycemic conditions and hence lower the risk of Type 2 Diabetes. An individual's diabetic diagnosis is based on the blood glucose parameters two hours after a meal of 140-199 mg/dL (Pre-Diabetic) and >200 mg/dL (Diabetic).[13] The presence of high lean body mass will not cause the blood glucose level to

plummet due to homeostatic responses where glycogen stored in muscle and liver cells is released into the bloodstream as glucose. This could pose a potential stone wall to individuals at risk; however other epigenetic responses could have been evoked as an effect of resistance training.

Theoretically, the physiological and metabolic strain could cause the skeletal muscles to undergo subjective changes in DNA methylation profiles. Thus, specific genes could govern the metabolism of lipids and glucose associated with Type 2 Diabetes, and the understanding would be that epigenetic control factors would regulate these genes' expression.

Role of Exercise Type and Intensity:

Participants	Type of exercise intervention and comparison group	Duration and intensity intervention	Tissue	Outcome measures		Level of evidence[a]
				epigenetic modification	outcome	
Healthy subjects						
Untrained patients n = 30 (20 F/10 M) Age: 46.1±8.56 y	Rehabilitation training No control group	5 days per wk for 2–3 h during 4 wk	Whole blood	Hypermethylation in one CpG site of *AMPKA2*	Improvement of metabolic responses to an energy imbalance	II-1
Subject with or without FH of T2D n = 28 M FH$^+$ 37.5±4 y FH$^-$ 37.5±5.2 y	Endurance training No control group	1 session of 1-h spinning class and 2 sessions of 1-h aerobic class per wk for 6 months	Skeletal muscle biopsy	Hypomethylation of genes involved in T2D development, respiratory chain, and calcium signaling pathways	Reduction of future risk of T2D	II-1
Healthy trained subjects n = 10 M Age: 30.5±5.5 y	Acute exercise test No control group	60 min of cycle ergometer exercise at 65% of peak watt before and after 12 wk of endurance training	Skeletal muscle biopsy	Upregulation of miR-1 and miR-133a	Improvement of insulin sensitivity by 19%	II-1
Healthy subjects n = 9 M Age: 23±5 y	Endurance training No control group	10 days: 45 min at 75% of VO_2max for 4 days 60 min at 75% of VO_2max one day 90 min at 75% of VO_2max one day 6 × 5 min intervals at 90–100% for 4 days	Skeletal muscle biopsy	Acute endurance exercise: Increase in miR-1, miR-133a, and miR-133b 3h after exercise and reduced miR-9, -23a, -23b and -31	Increase in transcription of genes involved in muscle regeneration and mitochondrial biogenesis	II-1
Healthy physically active subjects n = 12 M Age: 26.2±5.3 y	Acute exercise test No control group	45 min of one-legged knee extensor exercise at 60% of $watt_{max}$ before and after 7 days of bed rest	Skeletal muscle biopsy	Bed rest reduced miR-1 and miR-133a content; in addition, reduced HKII, SIRT1, HAD, and CS protein content	Bed rest induced whole body glucose intolerance and decreased muscle metabolic capacity as well as it abolished the exercise-induced adaptive gene response.	II-1
Healthy subjects n = 14 (9 M/5 F) Age 33±2 y	Acute exercise test No control group	48 min consisting of 4 sets, each composed of 8 min of cycling at 70%, 2 min at 90% of PHR, and 2 min rest	Skeletal muscle biopsy	Upregulation of miR-30, miR-128, and miR-378	Alteration of the skeletal muscle gene expression related to insulin sensitivity	II-1
Healthy subjects n = 10 M Age: 22.6±1.6 y	Acute exercise test No control group	Exercise test at 80% VO_2max for 15 min	Skeletal muscle biopsy	Increase in hydroxymethylation levels followed by hypomethylation of the *NR4A3* promoter	Improvement of metabolic function (normoglycemic levels) and endurance capacity	II-1
Two separate cohorts were used: 1. Healthy sedentary subjects n = 14 Age: 25±1 y 2. Healthy sedentary subjects n = 8 M	First cohorts group: Acute exercise test Second cohorts group: acute exercise test No control group	First cohorts group: Incremental exercise on cycle ergometer until the fatigue Second cohorts group: 2 isocaloric acute exercise trials at 40% and 80% VO_2max on at least 1 week of separate occasions	Skeletal muscle biopsy	First cohort: Intensity-dependent gene hypomethylation Second cohort: Hypomethylation of genes involved in T2D development	Beneficial effect on the prevention and treatment of type 2 diabetes	II-1
Healthy subjects n = 8 M Age: 21.1±2.2 y	Resistance exercise training No control group	Three sets of 8–12 repetitions during 8 wk at 80% of 1RM	Whole blood	Hypermethylation on 28,397 CpG site and hypomethylation on 28,987	Improvement of diabetes-related signaling pathways	II-1
Diseased subjects Obesity and T2D patients n = 17 (13 F/5 M) Age: 49±5 y	Resistance exercise training (n = 9) Endurance training (n = 8)	Resistance training group: 2–3 sets of 6–8 repetitions using machine weights during 16 wk Endurance training group: performed exercise on a cycle ergometer for 40–60 min during 16 wk	Skeletal muscle biopsy	Hypomethylation on inflammatory and glucose homeostasis genes	Better regulation of glycolytic flux	I

Figure 2: Exercise-induced epigenetic modifications related to insulin resistance and diabetes.[14]

Another study set out to understand the exercise-induced epigenetic modifications related to insulin resistance and diabetes and the results are presented in Figure 2. This study followed up on 50 individuals from each test group for 6 months.

Analyzing the data, the broad physical exercise spectrum ameliorates metabolic responses, insulin sensitivity, muscular regeneration, and Type 2 Diabetes risk. However, interesting results emerge based on the different types of training; rehabilitation, endurance, acute exercise, and resistance training. Outcomes directly related to the prevalence of Type 2 diabetes are tied to epigenetic alterations caused by acute exercise or resistance training groups. Furthermore, the study highlights the positive impact of resistance training on obese and Type 2 diabetic patients; improved glycolytic homeostasis induced by genes. Individuals following a strenuous and high-intensity exercise regimen reduce their risk of developing Type 2 diabetes and help them reduce the condition's effects. Studies highlight the high prevalence of sarcopenia in South-Indian populations, characterized by lower lean body mass and higher visceral fat percentages.[15]

Role of Abdominal Obesity:

Body composition synergistically works with physical training in playing a role for Type 2 Diabetes. Understanding that BMI has its limitations with unreliable measures of lean body mass, often waist circumference is the following optimal measurement procedure. Large waist circumference is often a result of excess subcutaneous fat, characterized by consuming a caloric surplus with minimal physical exertion. Visceral fat percentages often rise with this issue, and the scientific basis for the connection between abdominal obesity and diabetes is theoretically viable. Abdominal obesity enhances the risk of diabetes by a variety of released substances such as non-esterified fatty acids and adipocytokines such as tumor necrosis factor and decreased adiponectin.[16]

	CRP (mg/L)	P-value	IL-6 (pg/L)	P-value[b]
	Mean ± SE		Mean ± SE	
Body mass index[c]		0.269		0.666
Normal	3.73±1.22[e]		5.94±2.55[e]	
Overweight	4.54±0.66[e]		2.12±0.25[e]	
Obesity I	10.19±3.68[e]		3.96±1.02[e]	
Obesity II	10.09±1.85[e]		4.32±1.43[e]	
Obesity III	6.56±2.77[e]		6.35±2.55[e]	
Abdominal fat %[d]		0.023		0.002
Low	3.92 ± 0.86[e]		4.90± 2.26[ef]	
Average	5.53 ± 0.54[e]		2.68± 0.28[e]	
High	14.20 ± 1.24[f]		6.40± 2.08[f]	

Figure 3: Relation Between BMI and Abdominal Fat and Inflammatory Biomarkers in Patients with Type-2 Diabetes With Good Glycemic Control.[17]

	CRP (mg/L)		IL-6 (pg/L)	P-value[b]
			Mean ± SEM	
Body mass index[c]		0.02		0.047
Normal	5.42±2.08[e]		4.06±1.82[e]	
Overweight	6.79±1.18[e]		3.38±0.62[e]	
Obesity I	7.80±0.97[e]		3.89±0.34[e]	
Obesity II	11.17±1.16[f]		4.41±0.44[ef]	
Obesity III	16.66±3.78[g]		6.10±0.93[f]	
Abdominal fat %[d]		0.014		0.049
Low	8.82±4.25[e]		3.30±2.14[e]	
Average	8.72±0.71[e]		4.24±0.31[f]	
High	12.47±2.67[f]		4.42±0.60[f]	

Figure 4: Relation Between BMI And Abdominal Fat and Biomarkers of Inflammation in Patients with Type-2 Diabetes with Poor Glycemic Control.[17]

Figures 3 and 4 highlight how Type 2 Diabetes with Poor Glycemic Control has increased inflammation occurrences in pancreatic cells. This study was conducted using 198 individuals with Type 2 Diabetes Mellitus over the course of 3 months. Speculating could explain how epigenetic markers can be activated and repressed by DNA methylation and acetylation. The 'turning on and off' of epigenomic expression could play a role

DOI: 10.36838/v4i6.19

in inflaming associated cells where the inflammation disrupts pancreatic beta cells' sensitivity and response signals. Active control of nutritional caloric intake and physical exertion would create a metabolic deficit where the pancreatic alpha cells are primed to produce glucagon, chipping away at glycogen stores and depleting subcutaneous and visceral fat stores. Highlighting the data from Figure 3 shows how even in individuals with Good Glycemic Control, decreasing Abdominal Fat % from High to Low creates a decrease of 10.28 mg/L of C-Reactive Protein (CRP). CRP is a protein synthesized in the liver in response to inflammation and external cross-sectional studies have found a correlation ($P<0.01$) between CRP levels and central obesity. The study indicates that higher levels of adiposity in individuals lead to low-grade chronic inflammation which would hinder functional performance of smooth and skeletal muscles.[18] Poor muscular performance can hinder recovery and effectiveness of training programs which ultimately negate the target of reducing the effects of Type 2 Diabetes Mellitus. Although larger abdominal adiposity is generally explained by a sustained caloric surplus without adequate exercise, there may be a link between Type 2 Diabetes Mellitus and South Asia's consumption of the staple food; white rice. Referring to a study that followed 132, 373 participants for 9.5 years, the researchers set out to find a link between white rice consumption and incident diabetes by analyzing data from 21 countries. Results found a statistical association between higher white rice consumption (>450 g per day) and increased risk of Type 2 Diabetes (Hazard Risk of 1.20) within the whole study. However, whilst comparing different regions, it was observed that South Asia had the highest Hazard Risk of 1.61. This observation may be impacted by uncontrolled factors throughout the study's 9.5 years, however, comparing rice consumption may reveal new insights into the plausibility of rice consumption playing a significant role in South India's Diabetes prevalence. Data from the study show that China (HR of 0.78) had the second-highest rice consumption of 200 g/day whereas South India had the highest rice consumption of 630 g/day. This alarming disparity is supported by the higher risk of incident diabetes within the studied population and is an indicator to reduce white rice consumption to prevent the onset of diabetes.[19] Although external factors could have contributed to this statistical difference, loss of abdominal fat desaturates the circulatory stream from Low-Density Lipoproteins (LDLs) and lowers the occurrences of inflammatory responses.[20] If the buildup of LDLs continues to increase progressively, inflammation will give way to atherosclerotic progression as well. It is understood that inflammation caused by high abdominal fat percentage would be an early indicator of high-risk Type 2 Diabetes, especially amongst the largely sarcopenic South-Indian population.[21]

Comparison of Resistance Training Methods for Improved Glycemic Control:

A randomized controlled trial investigated different resistance training methods to improve glycemic control. The study followed 32 individuals through an 8-week training program. The individuals were randomly assigned to either an 8-week program of Hypertrophy Resistance Training (HRT) where exercises were performed with 2 sets at 10-12 repetitions or an 8-week program of Endurance Resistance Training (ERT) with 2 sets at 25-30 repetitions. The study's conclusions highlight HRT's specific maximal resistance training aids to significantly boost strength in target muscles compared to ERT. However, HRT showed no significant advantage to ERT in improving glycemic control where both groups exhibited comparable improvements.[22] As both resistance training methods provide comparable results, it would be beneficial to leave the option to the patient's discretion and preference.

Conclusion

This study has employed theorized connections between exercise and epigenetic control factors and cross-referenced studies to establish a ground understanding of their influence on Type 2 diabetes within Tamil Nadu, India. It has come to light that occupational activity levels, rehabilitation, and endurance training are insufficient in combatting the epidemic of Type 2 diabetes. Such altercations pose morbid threats to global health systems and require awareness and appropriate action. Evidence suggests the positive effect of regular resistance training and other high-intensity exercise plans on epigenetic alterations in physiological and metabolic systems. It has also been shown that there is no baseline difference between hypertrophy resistance training and endurance resistance training, leaving the choice up to the preference of the patient. Furthermore, the public's use of resistance training could also be woven into their daily lifestyles. Commuting using bicycles and voluntarily opting to use staircases are subtle strategies the average individual could employ to resistance train. The study also highlights a concern with South India's alarming consumption of white rice; a staple food. Higher white rice consumption is linked to a higher risk of developing Type 2 Diabetes Mellitus; hence, it is suggested to reduce rice consumption as a preemptive measure to reduce the risk of Type 2 Diabetes. As such, global awareness would need to be raised, emphasizing resistance training and updating the public's understanding of 'healthy' lifestyles.

An extension to this study can be conducted to quantifiably measure the impact of physical exercise on the epigenetic control of Type 2 diabetes in Tamil Nadu, India. Taking a group of consensual diabetic volunteers and a healthy control volunteer, a 6-month long study could provide different training method regimens at different intensities. In addition, at regular intervals, a glycated hemoglobin (A1C) test could be conducted to measure the average blood sugar level for the past three months.[23] The quantitative analysis could uncover new insights into practical strategies that already affected individuals could employ to alleviate their symptoms and potentially even reverse the effects of Type 2 Diabetes.

Acknowledgments

The author would like to extend sincere thanks to Dr. Satendra Kumar Multani (MBBS, MD, DNB, MNAMS, DM(Endo), FACE, MRCP-SCE (Endo)-UK).

References

1. Tripathy, J.; Thakur, J.; Jeet, G.; Chawla, S.; Jain, S.; Pal, A.; Prasad, R.; Saran, R. Prevalence And Risk Factors Of Diabetes In A Large Community-Based Study In North India: Results From A STEPS

Survey In Punjab, India. Diabetology & Metabolic Syndrome 2017, 9 (1).
2. Dansinger M. Type 2 Diabetes, 2020. WebMD. https://www.webmd.com/diabetes/type-2-diabetes (accessed Aug 5, 2021).
3. U.S National Library of Medicine. Type 1 diabetes. https://medlineplus.gov/genetics/condition/type-1-diabetes/#causes (accessed Aug 10, 2021).
4. B, Cornell, B. DNA Structure, 2016. BioNinja. https://ib.bioninja.com.au/standard-level/topic-2-molecular-biology/27-dna-replication-transcri/ (accessed Sep 13, 2021).
5. Curradi, M.; Izzo, A.; Badaracco, G.; Landsberger, N. Molecular Mechanisms Of Gene Silencing Mediated By DNA Methylation, 2002. Molecular and Cellular Biology 2002, 22, 3157-3173.
6. Cornell, B. Gene Expression, 2016. BioNinja. https:// ib.bioninja.com.au/higher-level/topic-7-nucleic-acids/72-transcrip tion-and-gene/gene-expression.html (accessed Aug 10, 2021).
7. Census India. DECADAL VARIATION IN POPULATION SINCE 1901; New Delhi, 2011; p. 1.
8. Gupta, S.; Singh, Z.; Purty, A.; Kar, M.; Vedapriya, D.; Mahajan, P.; Cherian, J. Diabetes Prevalence And Its Risk Factors In Rural Area Of Tamil Nadu. Indian Journal of Community Medicine 2010, 35, 396-399.
9. ICMR-National Institute of Nutrition. Nutrient Requirements For Indians; Hyderabad, 2020.
10. Heslehurst, N.; Sattar, N.; Rajasingam, D.; Wilkinson, J.; Summerbell, C.; Rankin, J. Existing Maternal Obesity Guidelines May Increase Inequalities Between Ethnic Groups: A National Epidemiological Study Of 502,474 Births In England. BMC Pregnancy and Childbirth 2012, 12.
11. Oommen, A.; Kapoor, N.; Thomas, N.; George, K. Prevalence And Clinical Characteristics Of Individuals With Newly Detected Lean Diabetes In Tamil Nadu, South India: A Community-Based Cross-Sectional Study. International Journal of Diabetes in Developing Countries 2019, 39, 680-684.
12. Steeves, J.; Tudor-Locke, C.; Murphy, R.; King, G.; Fitzhugh, E.; Harris, T. Classification Of Occupational Activity Categories Using Accelerometry: NHANES 2003–2004. International Journal of Behavioral Nutrition and Physical Activity 2015, 12.
13. Diabetes - Diagnosis and treatment - Mayo Clinic https://www.mayoclinic.org/diseases-conditions/diabetes/diagnosis-treatment/drc-20371451 (accessed Oct 19, 2021).
14. Barrón-Cabrera, E.; Ramos-Lopez, O.; González-Becerra, K.; Riezu-Boj, J.; Milagro, F.; Martínez-López, E.; Martínez, J. Epigenetic Modifications As Outcomes Of Exercise Interventions Related To Specific Metabolic Alterations: A Systematic Review. Lifestyle Genomics 2019, 12, 25-44.
15. Sreepriya, P.; Pillai, S.; Nair, n.; Rahul, A.; Pillai, S.; Nair, A. Prevalence And Associated Factors Of Sarcopenia Among Patients Underwent Abdominal CT Scan In Tertiary Care Hospital Of South India. Journal of Frailty, Sarcopenia and Falls 2020, 05, 79-85.
16. Freemantle, N.; Holmes, J.; Hockey, A.; Kumar, S. How Strong Is The Association Between Abdominal Obesity And The Incidence Of Type 2 Diabetes?. International Journal of Clinical Practice 2008, 62, 1391-1396.
17. Bawadi, H.; Katkhouda, R.; Tayyem, R.; Kerkadi, A.; Bou Raad, S.; Subih, H. <P>Abdominal Fat Is Directly Associated With Inflammation In Persons With Type-2 Diabetes Regardless Of Glycemic Control – A Jordanian Study</P>. Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy 2019, Volume 12, 2411-2417.
18. Forouhi, N.; Sattar, N.; McKeigue, P. Relation Of C-Reactive Protein To Body Fat Distribution And Features Of The Metabolic Syndrome In Europeans And South Asians. International Journal of Obesity 2001, 25 (9), 1327-1331.
19. Bhavadharini, B.; Mohan, V.; Dehghan, M.; Rangarajan, S.; Swaminathan, S.; Rosengren, A.; Wielgosz, A.; Avezum, A.; Lopez-Jaramillo, P.; Lanas, F.; Dans, A.; Yeates, K.; Poirier, P.; Chifamba, J.: Alhabib, K.; Mohammadifard, N.; Zatońska, K.; Khatib, R.; Vural Keskinler, M.; Wei, L.; Wang, C.; Liu, X.; Iqbal, R.; Yusuf, R.; Wentzel-Viljoen, E.; Yusufali, A.; Diaz, R.; Keat, N.; Lakshmi, P.; Ismail, N; Gupta, R.; Palileo-Villanueva, L.; Sheridan, P.; Mente, A.; Yusuf, S. White Rice Intake And Incident Diabetes: A Study Of 132,373 Participants In 21 Countries. Diabetes Care 2020, 43 (11), 2643-2650.
20. Shay Martin, S. Cholesterol in the Blood https://www.hopkinsmedicine.org/health/conditions-and-diseases/high-cholesterol/cholesterol-in-the-blood#:~:text=Cholesterol%20and%20other%20fats%20are,%2Ddensity%20lipoproteins%20(HDL) (accessed Oct 19, 2021).
21. Jukema, R.; Ahmed, T.; Tardif, J. Does Low-Density Lipoprotein Cholesterol Induce Inflammation? If So, Does It Matter? Current Insights And Future Perspectives For Novel Therapies. BMC Medicine 2019, 17.
22. Egger, A.; Niederseer, D.; Diem, G.; Finkenzeller, T.; Ledl-Kurkowski, E.; Forstner, R.; Pirich, C.; Patsch, W.; Weitgasser, R.; Niebauer, J. Different Types Of Resistance Training In Type 2 Diabetes Mellitus: Effects On Glycaemic Control, Muscle Mass And Strength. European Journal of Preventive Cardiology 2012, 20 (6), 1051-1060.
23. Type 2 diabetes - Diagnosis and treatment - Mayo Clinic https://www.mayoclinic.org/diseases-conditions/type-2-diabetes/diagnosis-treatment/drc-20351199 (accessed Aug 11, 2021).

Author

Jerosh Antoine Susaiprakash is a year 13 student in the International Baccalaureate Diploma Programme at Emirates International School Jumeirah. Jerosh is studying Biology, Chemistry, and Mathematics: Applications and Interpretations at Higher Level.

DOI: 10.36838/v4i6.19

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

CO_2 Storage Urease Enzyme-Mediated Calcite Precipitation: Implication for CO_2 Leakage Prevention

Jenny Yu, Sarah Kim
Portsmouth Abbey School, 285 Corys Lane, Portsmouth, RI, 02871, USA; yubeen.lee@stpaulseoul.org

ABSTRACT: The increase in anthropogenic CO_2 emissions has been acknowledged as an important issue arising within our environment. In response, the Korean government announced its goals to reduce CO_2 emissions to 40% of the current level by 2030. One of the possible techniques for reducing CO_2 is geologic CO_2 sequestration also termed carbon capture and storage. This is a process aimed at mitigating the release of CO2 into the atmosphere to reduce the contribution of fossil fuel emissions to global warming and ocean acidification. The storage formation ideally should have a structure with high porosity and permeability for effective CO_2 injection. However, a layer of low permeability (cap rock) with gas-tight installations and cementing materials must be used to seal the formation to prevent CO_2 leakage. Recently, the enzyme-induced carbonate precipitation (EICP) technique has been considered as one of the solutions: ureolysis-driven $CaCO_3$ precipitation was proposed and studied to be applied *in-situ* as a sealant for treating fractures and high permeability areas in cap rocks and bore cement. In this study, calcite precipitation experiments were conducted to observe the effect of enzyme amounts on the EICP. The results show that the amount of calcite formed by the EICP increased with the increase in enzyme amounts. Calcite precipitation was visible to the naked eye within a day and relatively more rapid calcite precipitation was observed as the enzyme amounts increased. This study shows that the amount and rate of calcite precipitation increased in proportion to the enzyme amounts.

Introduction

The increase in anthropogenic CO_2 emissions is acknowledged as an important issue arising within our environment. In the case of Korea, the government has announced its goal to reduce CO_2 emissions by 2030 to 40% of 2018 levels and achieve net-zero emissions by 2050.[1] Various technologies for reducing emissions of CO_2 to mitigate climate change have attracted the interest of scientists. One of the techniques is to capture and store carbon dioxide in a deep geologic formation.

The mineral carbonation of CO_2, converting atmospheric CO_2 into carbonate minerals, has been presented as a geologically stable and environmentally safe way to store carbon dioxide.[2] The major issue with spontaneous carbonate mineral formation is that this reaction tends to have slow reaction rates and is highly dependent on pH.[3] Biological carbonate mineral formation, called microbially-induced carbonate precipitation (MICP), has recently been suggested to participate in CO_2 sequestration.[4] This biological process is commonly involved in the hydrolysis of urea by the urease-producing bacteria.

On the other hand, the increasing demand for cement brings along environmental concerns because cement production is a large source of carbon dioxide emissions. There has been a large amount of research to find substitutes for cement that utilizes more eco-friendly sources. Biocementation is one method presented as a substitute, which includes the sustainable use of free urease enzymes. Enzyme-Induced *Carbonate Precipitation* (EICP) is the process of precipitating calcium carbonate utilizing a free urease enzyme, one of the innovative engineering uses of urease. Because no living organism is actively engaged in the precipitation process, EICP is considered a biomimetic method. The calcium carbonate precipitate formed from this process can effectively sequester and reduce the leakage of CO_2 from storage formations by plugging pores.

Urease-aided $CaCO_3$ mineralization, or urinalysis-driven $CaCO_3$ mineralization, depends on the activity of urease. The benefit of this process is that, by reproducing $CaCO_3$ bio-formation in nature, the process can be performed in an environmentally benign manner and, importantly, in field applications. Accordingly, the process offers exciting innovative potential in multiple engineering fields such as geotechnical engineering, bioremediation for heavy metals, plugging oil reservoir bedrocks for the enhancement of oil recovery, and reducing leakage in geologic CO_2 sequestration.[5,7] Two methods of ureolysis-driven calcite precipitation are summarized in Table 1.[8]

Table 1: The types of biomineralization of calcium carbonate

Type	Microbially-Induced Carbonate Precipitation (MICP)	Enzyme-Induced Carbonate Precipitation (EICP)
Condition	microbial organism	a free urease enzyme
Source	soil bacterium	extracted from plants (commonly jack beans)
Advantage	- widespread usage - unaffected by temperature	- commercial availability - smaller molecular size - efficient penetration of urease - natural degradation

DOI: 10.36838/v4i6.20

Disadvantages	- expensive (bacteria) - potential cause of bioplugging - lower durability of formed precipitate	restricted usage

Based on previous research on Enzyme-Induced Carbonate Precipitation (EICP), the objectives of this study are 1) examining the effect of the concentration of the enzyme in a solution with constant concentrations of urea and calcium chloride and 2) understanding calcite precipitation rate and morphology of the precipitates in a laboratory environment. These experimental data can be used as basic data for applying EICP to the reduction of CO_2 leakage in geologic CO_2 sequestration.

Theoretical Background (Urease-aided calcium carbonate mineralization):

The enzyme-induced carbonate precipitation (EICP) is triggered by the catalytic action of urease in the hydrolysis of urea. The products of the reaction are carbonic acid and ammonia (1):

$$H_2N\text{-}CO\text{-}NH_2 + 2H_2O \xrightarrow{\text{Urease}} H_2CO_3 + 2NH_3 \quad (1)$$

The products equilibrate in water to give bicarbonate, ammonium, and hydroxide ions, respectively:

$$H_2CO_3 \leftrightarrows HCO_3^- + H^+ \quad (2)$$

$$NH_3^+ H_2O \leftrightarrows NH_4^+ + OH^- \quad (3)$$

The production of hydroxide ions from reaction (3) brings about an increase in pH, which in turn leads to the formation of carbonate ions:

$$HCO_3^- + OH^- \leftrightarrows CO_3^{2-} + H_2O \quad (4)$$

In the presence of dissolved Ca^{2+}, the ions combine and calcium carbonate precipitates:

$$Ca^{2+} + CO_3^{2-} \rightarrow CaCO_3(s) \quad (5)$$

The overall process can thus be presented:

$$H_2N\text{-}CO\text{-}NH_2 + 2H_2O + Ca^{2+} \xrightarrow{\text{Urease}} 2NH_4^+ + CaCO_3(s) \quad (6)$$

As shown in the precipitation process, 1 mole of urea produces 1 mole of calcite.

Methods

For this study, two types of solutions were prepared:

1) a urease solution using a urease enzyme (U1500, Sigma Aldrich) and

2) an equimolar urea-$CaCl_2$ solution using urea (U5378, Sigma Aldrich, St. Louis, Mo, USA) and $CaCl_2$ (C3881, Sigma Aldrich).

The urease solutions were prepared at concentrations of 1000 µL, 2000 µL, and 3000 µL by dissolving the urease enzyme into distilled water, while the urea-$CaCl_2$ solutions were prepared at concentrations of 0.5 M. To identify the effect of concentrations of the enzyme on the production of calcite, each urease solution was mixed with 0.5 M urea-$CaCl_2$ solutions to precipitate calcite.

In this experiment, 20 mL of urea-$CaCl_2$ solution and 20 mL of urease solution were mixed in conical tubes to make a total solution volume of 40 ml. A total of 20 samples were prepared for the batch experiments and one blank sample without the enzyme was also tested. Those samples were reacted at 30 °C in a water bath without agitation. During the experiment, the mass of precipitated calcite was measured from a selected tube at 1 h, 3h, 5 h, 8 h, 15 h, 24 h, 48 h, and 72 h. The tubes were centrifuged at 3000 rpm for 10 minutes to separate the remaining liquid and precipitate. After removing the supernatant, the tubes were dried in an oven at 50°C for 48 hours, and the amount of precipitated calcium carbonate was determined by measuring the change in the mass of the tubes before and after the experiment. To identify the precipitated mineral, the precipitates were recovered and analyzed by scanning electron microscopy with energy-dispersive spectroscopy **(SEM-EDS)**.

The process of this experiment can be summarized in Figure 1.

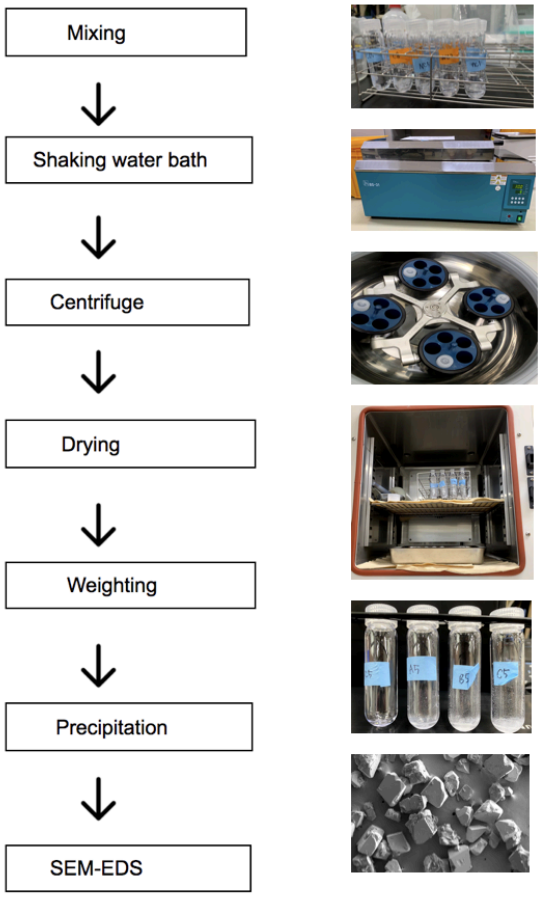


Figure 1: Schematic procedure and figures of EICP experiment. From these experiments, we determined the calcite amounts and observed morphology.

Results and Discussion

1) The change in the amount of calcite precipitation according to enzyme concentration:

Based on the method above, the results of measuring the mass of the blank bottle and calcite precipitated bottle according to three different enzyme concentrations are shown in Table 2. As the amount of enzyme increases, precipitation

DOI: 10.36838/v4i6.20

rates increase, the amount of calcite precipitated increases in proportion. One hour after the experiment, 0.0011 g was precipitated at 1000 μL enzyme concentration (hereinafter, concentration A), while 0.0151 g was precipitated at 3000 μL enzyme concentration (hereinafter, concentration C). This shows that the C concentration is about 1,372-fold higher than the A concentration. This tendency shows a similar trend in the 24 hours (546-fold) and 72 hours (386-fold) experimental results. The maximum difference is 1,728-fold, which is after 15 hours, and the minimum difference is 386-fold, which is after 72 hours (Figure 2).

Table 2: Experiment results of calcite precipitation depending on the amounts of enzyme concentrations (unit= g)

	A (1000 μL)			B (2000 μL)			C (3000 μL)		
	Blank*	B+P**	Net value	Blank	B+P	Net value	Blank	B+P	Net value
1hr	21.5658	21.5669	**0.0011**	21.6865	21.7001	**0.0136**	21.6855	21.7006	**0.0151**
3hr	21.8677	21.8708	**0.0031**	22.0082	22.0379	**0.0297**	21.891	21.9297	**0.0387**
5hr	21.6681	21.6834	**0.0153**	21.5147	21.5801	**0.0654**	21.8089	21.9001	**0.0912**
8hr	22.0029	22.0358	**0.0329**	21.5273	21.6884	**0.1611**	22.0003	22.1912	**0.1909**
15hr	22.0383	22.0490	**0.0107**	21.6929	21.8675	**0.1746**	21.6967	21.8817	**0.1850**
24hr	21.6924	21.7772	**0.0848**	21.8385	22.1581	**0.3196**	21.8771	22.3402	**0.4631**
48hr	21.5614	21.6999	**0.1385**	21.6905	22.1082	**0.4177**	21.6527	22.2484	**0.5957**
72hr	21.6790	21.8712	**0.1922**	22.0613	22.5812	**0.5199**	21.8932	22.6365	**0.7433**

*blank bottle
**calcite precipitation bottle

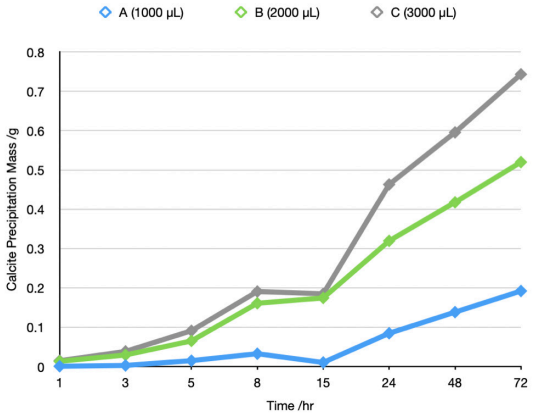


Figure 2: Changes in calcite precipitation according to three different enzyme concentrations.

The precipitate is white and exists as suspended particles in water. Calcite precipitation was visible to the naked eye within a day: 24 hours for concentration A, 15 hours for concentration B, and 8 hours for concentration C (Figure 3). Calcite precipitation occurred more rapidly as the enzyme concentration increased.

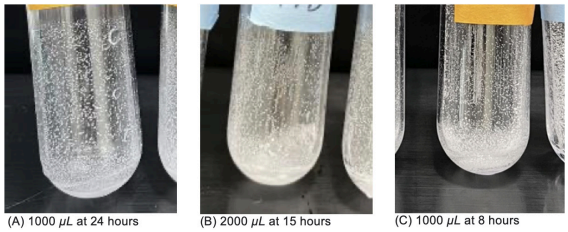


Figure 3: Calcite precipitation depending on the concentrations of enzymes.

2) The change in the amount of calcite precipitation according to reaction time:

According to reaction equation 6, 0.5M CaCl2 -urea solution must produce 0.5M calcite under equilibrium conditions. However, in the equilibrium state of the current experiment, the theoretical precipitation amount of 0.5M calcite is 2g. After 72 hours, the maximum precipitation at concentration C was 0.7433g, which is about 37% of the theoretical value.

Moreover, the three different enzyme concentrations show different trends over time. Based on the test results, concentration A showed the lowest precipitation rate whereas concentrations B and C showed relatively rare. At concentration A, the change in mass of precipitation is 0.01 g from 0 to 15 hours, 0.08 g from 15 to 24 hours, and 0.13 g from 24 to 48 hours. The difference between the change in precipitation mass in the 0 to 15 h range and the 24 to 48 h range is about 8 times. From the 15 to 24 h range to the 24 to 48 h range, the precipitation rate decreases rapidly by about 1.6 times. However, for concentration C, the precipitation mass increased by 0.18 g from 0 to 15 hours, 0.46 g from 15 to 24 hours, and 0.59 g from 24 to 48 hours. The change in mass during the 15 to 24 h increased 2.5 times faster than the 0 to 15hr range, while the mass change within the 24 to 48 hr. range was 1.3 times faster than the 15 to 24 hr. range. This shows a slight decrease in the rate of precipitation as the reaction proceeded. Nonetheless, the experimental results show that the amount and rate of calcite precipitation increases according to the increase in the enzyme concentration. This suggests that the concentration of the enzyme is not enough to completely hydrolyze the 0.5 M urea in this study.

In a previous study, it was also found that an increase in urease concentration enhanced the rate of $CaCO_3$ precipitation.[12]

3) Morphology of calcite:

Figure 4 shows the morphology of calcite after 72 hours of precipitation under the condition of concentrations A, B, and C respectively. Noticeable differences in the calcite morphology due to the change in concentration were not observed within this experiment. For further applications of calcite outside of the experiment setting, more experiments are needed to test the stability and persistence based on the size and morphology of calcite when filling the crack of subsurfaces.

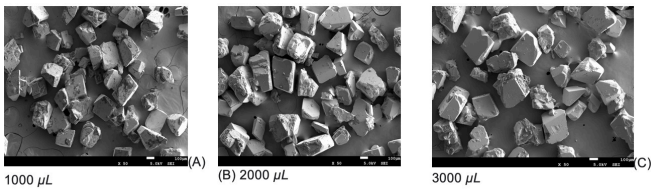


Figure 4: SEM-EDS Image of calcite after 72 hours of precipitation for different enzyme concentrations.

4) Implications for future experiments:

Factors that influence the precipitation process primarily include the concentration of urease and the concentrations of urea and Ca^{2+}, in addition to environmental factors such as temperature and salinity.[11,13]

In this regard, for a given concentration of urease, there would be optimal urea and Ca^{2+} concentrations. Yet, these

benefits are counterbalanced by the detrimental effects including the accumulation of salts and urea in the material. The temperature was also found to affect the rate of $CaCO_3$ precipitation in the enzymatic process, and practically not influence the bacterial process.[13]

Regardless of using MICP or EICP, the ureolysis-driven calcite precipitation produces the same by-products. One of which is the ammonium ion NH^{4+} resulting from the hydrolysis of urea (Equations. (3) and (6)). The ion poses an ecological risk that includes NH_3 toxicity to humans and acidification arising from NH_3 nitrification, in addition to stone discoloration.[14] These effects, however, can be minimized by using low and balanced urea concentrations or exchanged zeolite.[15]

Conclusion

In our experiment study, calcite precipitation experiments were conducted to observe the effect of enzyme amounts on the EICP.

1) The amount of calcite formed by the EICP increased with the increase in enzyme concentrations. Concentration C (3000 μL) resulted in a 550 times greater amount of calcite formation compared to concentration A (1000 μL) at 24 hours and about 387 times greater at 72 hours.

2) Calcite precipitation was white-colored, suspended particles, clearly visible to the naked eye within a day. Faster calcite precipitation was observed as the enzyme amounts increased.

3) There were no significant differences observed in the calcite morphology due to the change in enzyme concentration.

This study showed that the amount and rate of calcite precipitation increased in proportion to the enzyme amounts. Further studies need to identify the effects of other factors such as urea and calcium concentrations, temperature, and salinity on EICP to find optimal enzyme concentrations for practical applications of the technique.

References

1. Kim E-Y and Kim H., Gov't raises goal of cutting CO2 emissions to 40% by 2030., https://www.korea.net/NewsFocus/policies/ view?articleId=205222, 2021.
2. Ramanan, R., Kannan, K., Sivanesan, S., Mudliar, S., Kaur, S., Tripathi, A., et al., 2009. Biosequestration of carbon dioxide using carbonic anhydrase enzyme purified from Citrobacter freundii. World J. Microbiol. Biotechnol. 25, 981–987.
3. Zhu, X., Logan, B.E., 2014. Microbial electrolysis desalination and chemical-production cell for CO2 sequestration. Bioresour. Technol. 159, 24–29.
4. Okwadha GDO, Li J. Optimum conditions for microbial carbonate precipitation. Chemosphere 2010; 81:1143–8.
5. Anbu P, Kang C-H, Shin Y-J, So S-S. Formation of calcium carbonate minerals by bacteria and its multiple applications. Springer Plus 2016; 5:250
6. DeJong JT, Mortensen BM, Martinez BC, Nelson DC. Bio-mediated soil improvement. Ecol En2010; 36:197–210.
7. Kumari D, Qian X-Y, Pan X, Achal V, Li Q, Gadd GM. Microbially-induced carbonate precipitation for immobilization of toxic metals. Adv Appl Microbiol 2016; 94:79–108.
8. Krajewska B, Urease-aided calcium carbonate mineralization for engineering applications: A Review. J Adv Res 2018, 13, 59-67
9. Phillips AJ, Gerlach R, Lauchnor E, Mitchell AC, Cunningham AB, Spangler L. Engineered applications of ureolytic biomineralization: a review. Biofouling 2013; 29:715–33.
10. Okwadha GDO, Li J. Optimum conditions for microbial carbon ate precipitation. Chemosphere 2010; 81:1143–8.
11. De Muynck W, Verbeken K, De Belie N, Verstraete W. Influence of urea and calcium dosage on the effectiveness of bacterially indu ced carbonate precipitation on limestone. Ecol Eng 2010; 36:99–111.
12. Nemati M, Voordouw G. Modification of porous media perme ability, using calcium carbonate produced enzymatically in situ. En zyme Microb Tech 2003; 33:635–42.
13. Nemati M, Greene EA, Voordouw G. Permeability profile modifi cation using bacterially formed calcium carbonate; comparison with enzymic option. Process Biochem 2005; 40:925–33.
14. Achal V, Mukherjee A, Kumari D, Zhang Q. Biomineralization for sustainable construction – A review of processes. Earth-Sci Rev 2015; 148:1–17.
15. Lee S-H, Kim J-M. A study on Enzyme-induced carbonate pre cipitation Technique for Soil reinforcement without Releasing an Environmental Contaminant. Korea Geotech. Soci., V37, 2021, 39-47.

Authors

Jenny Yu is a senior at Portsmouth Abbey School in Portsmouth, RI. She enjoys studying chemistry and environmental science. She is interested in continuing her renewable energy research in college.

Sarah Kim attends Chadwick International in South Korea as a senior. Her deep-rooted interests in environmental science and engineering have been greatly reflected in this project and are aspiring to continue to extend her knowledge in this field through college research.

DOI: 10.36838/v4i6.20

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Meteorological Factors Associated with Elevated Levels of Daily $PM_{2.5}$ Concentrations in Seoul, South Korea, in 2019

Yubeen Lee
Saint Paul Preparatory Seoul, 14-8, Seocho Jungang-ro 31-gil, Seocho-gu, Seoul, 06593, Korea; yubeen.lee@stpaulseoul.org

ABSTRACT: Studies have revealed the adverse health impact of particulate matter 2.5 ($PM_{2.5}$). Elevated $PM_{2.5}$ levels are often observed in cities with abundant sources of emissions of pollutants. In addition to emissions, $PM_{2.5}$ concentrations can vary with meteorological conditions. This study aimed to identify meteorological variables associated with high levels of $PM_{2.5}$ in Seoul, South Korea. An analysis was conducted using data associated with daily meteorological variables in Seoul from 2019, downloaded from the Korea Meteorological Administration (KMA) and Air Korea webpages. These variables included daily precipitation duration and amount, maximum and average wind speed, dominant and prevailing wind direction, relative humidity, vapor pressure, and temperature. The relationship between these variables and $PM_{2.5}$ concentrations was analyzed using two datasets: one containing data from the entire year of 2019 and the other from January, February, and March when $PM_{2.5}$ levels were relatively high. Daily average $PM_{2.5}$ concentrations were categorized into two groups: acceptable ($PM_{2.5} \leq 25$ μg/m³) and poor ($PM_{2.5} > 25$ μg/m³) air quality. When analyzing data from the entire year, it was observed that the median $PM_{2.5}$ concentration was statistically higher when westerly winds were observed [Probability value (P-value) <0.001]. By contrast, the three-month sub-analysis did not show any statistically significant differences in median $PM_{2.5}$ levels between westerly and non-westerly winds (P =0.390 for dominant wind direction, P =0.846 for prevailing wind direction). Although average wind speed, vapor pressure, and a prevailing westerly wind direction showed a significant correlation with $PM_{2.5}$ levels in data for the whole year, interpretations of the link between these meteorological variables and high levels of $PM_{2.5}$ were limited, partly due to seasonal variation. Further sub-analysis of the three-month dataset, corresponding with winter and early spring, showed that maximum wind speed and vapor pressure are significant factors associated with poor air quality. The results of this study can be used to predict $PM_{2.5}$ levels in Seoul, South Korea.

KEYWORDS: Earth and Environmental Sciences; Atmosphere Science; $PM_{2.5}$; Meteorological Conditions; Air Quality.

Introduction

Elevated levels of particulate matter (PM) degrade visibility and threaten public health.[1-3] Concentrations of PM are affected by various factors, including emissions of air pollutants, meteorological conditions, and wind-transported air pollution from foreign countries.[4,5] Recently, emissions in Seoul, South Korea, have slowly decreased; thus, polluted air from foreign sources has gained increasing attention.[6-8] Investigators have asserted that air pollutants are transported from China by westerly winds, particularly during winter and early spring, resulting in frequent and severe haze episodes in western South Korea and other East Asian countries.[9,10] Previous studies have analyzed ground-level PM concentrations using satellite-based data,[11] satellite observations, and process-based models,[12] and conducted simulations using a regional air quality modeling system.[1] A large number of studies recently investigated the impact of meteorological conditions such as surface temperature, pressure, humidity, and wind speed on PM variability.[1,13] Among different PMs, $PM_{2.5}$ is known to be especially harmful to the human body; because of its smaller size compared to PM_{10}, it can infiltrate the lungs and bloodstream when breathing.[14] Therefore, this study aimed to investigate the meteorological factors related to high levels of $PM_{2.5}$ in Seoul.

Methods

Daily meteorological and air pollution data in Seoul from 2019 were downloaded from the Korea Meteorological Administration (KMA) (https://data.kma.go.kr) and Air Korea webpages (https://www.airkorea.or.kr).[15,16] The variables analyzed included daily precipitation duration and amount, maximum and average wind speed, dominant and prevailing wind direction, relative humidity, vapor pressure, and temperature. Dominant wind direction was defined as the direction of the wind with the highest speed over a particular point, and prevailing wind direction was defined as the direction that the wind blows most often in a location. Wind direction was categorized into eight directions: north, northeast, east, southeast, south, southwest, west, and northwest. The daily average $PM_{2.5}$ concentration in Seoul was categorized into two groups: acceptable air quality ($PM_{2.5} \leq 25$ μg/m3) and poor air quality ($PM_{2.5} > 25$ μg/m³). The threshold of 25 μg/m³ was selected instead of 35 μg/m³, which is the daily air quality standard for $PM_{2.5}$ in Korea because exposure to $PM_{2.5} > 25$ μg/m3 is known to be harmful to sensitive groups, including people with heart or lung disease, older adults, children, and teenagers.[17]

The statistical analysis was conducted using two datasets: one containing the annual data for 2019 and the other

DOI: 10.36838/v4i6.21

comprising data for January, February, and March, when relatively high levels of $PM_{2.5}$ were observed. The three-month sub-analysis was performed to minimize the effects of seasonal variations in data. Levels of $PM_{2.5}$ were analyzed in relation to eight dominant and prevailing wind directions using a Kruskal-Wallis test, which is used to find the difference between medians of three or more groups.[18] *Post-hoc* analysis was performed for multiple comparisons using the Dunnett C test, often used to compare one group with others.[19] Furthermore, the difference in $PM_{2.5}$ levels was compared between the westerly and non-westerly dominant and prevailing wind directions to specifically investigate the effect of westerly winds on $PM_{2.5}$ levels using Wilcoxon rank-sum test.

Univariate and multivariate analyses were performed to identify the meteorological factors associated with high levels of air pollution in Seoul. Meteorological variables were compared between the acceptable and poor air quality groups using Fisher's exact test for categorical variables and the Wilcoxon rank-sum test for continuous variables.[20,21] Variables with probability values (P-values)< 0.1 obtained through univariate analysis were selected to perform the multivariate logistic regression analysis,[22] for which a forward stepwise selection mode was used; this mode allowed an iterative entry of variables based on test score P-values< 0.05 and achieved the removal of variables with a probability of 0.10. All statistical analyses were performed using the SPSS version 23.0 software (IBM Corp., Armonk, NY, USA). Statistical significance was set at P< 0.05.

Results and Discussion

The monthly average $PM_{2.5}$ and PM^{10} concentrations showed an apparent seasonal variation in particle levels. Particle concentrations were relatively high in winter and early spring (January, February, and March) (Figure 1). Thus, additional analyses were conducted for the three-month period.

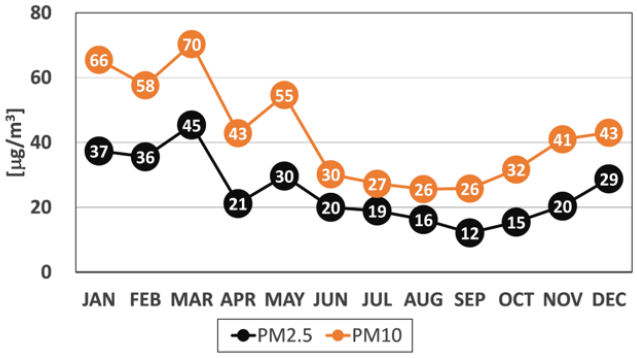


Figure 1: Monthly average $PM_{2.5}$ concentrations in Seoul, Korea, in 2019. Particle concentrations were relatively high in winter and early spring (January, February, and March).

Differences in $PM_{2.5}$ Levels between Wind Directions:

The median $PM_{2.5}$ concentrations of the eight wind directions were compared (Table 1). The Kruskal-Wallis test showed that $PM_{2.5}$ levels varied significantly between different dominant wind directions (P< 0.001). Multiple further post hoc comparisons were made using the Dunnett C test. The results showed that the differences were particularly significant between northerly and southerly, northerly and northwesterly, and easterly and westerly winds (P< 0.05). An identical analysis was conducted for the difference in $PM_{2.5}$ levels between prevailing wind directions using the Kruskal-Wallis test, the results of which also showed that $PM_{2.5}$ levels vary significantly between different prevailing wind directions (P< 0.001). Dunnett's *post hoc* multiple comparisons showed that the difference in $PM_{2.5}$ was significant between northerly and westerly winds as well as southeasterly and westerly winds (P< 0.001).

Table 1: Median $PM_{2.5}$ levels in relation to wind direction in 2019. Westerly winds were observed most frequently and associated with higher $PM_{2.5}$.

Wind direction	Dominant wind direction		Prevailing wind direction	
	No.	$PM_{2.5}$ [μg/m³] [Median (25th, 75th)]	No.	$PM_{2.5}$ [μg/m³] [Median (25th, 75th)]
N	37	14.0 (6.6, 19.6)	47	12.8 (7.2, 21.5)
NE	4	34.2 (28.6, 37.4)	35	19.7 (15.9, 27.0)
E	41	18.7 (10.8, 25.0)	46	19.5 (11.5, 27.4)
SE	5	22.6 (13.53, 24.0)	2	13.9 (13.7, 14.2)
S	49	20.3 (16.1, 28.0)	21	20.4 (13.7, 25.2)
SW	17	20.4 (12.9, 25.2)	18	17.6 (12.0, 26.1)
W	161	23.4 (16.9, 35.5)	180	23.4 (17.5, 34.1)
NW	51	21.1 (16.2, 31.0)	16	21.0 (15.1, 25.9)

Among the eight wind types, westerly winds were observed most frequently (Table 1), during which $PM_{2.5}$ levels were relatively high. Thus, further analysis was conducted to compare the differences in $PM_{2.5}$ levels between westerly and non-westerly winds. When analyzing the 2019 dataset, the median $PM_{2.5}$ level was significantly higher when westerly winds were observed (P< 0.001) (Table 2). However, when analyzing the January–March dataset, the difference in median $PM_{2.5}$ levels between wind directions was not statistically significant (P= 0.390 for dominant wind direction, P= 0.846 for prevailing wind direction). The results indicated that in addition to meteorological variables, other variables could affect $PM_{2.5}$ levels in Seoul.[4,5] Given that emissions emitted from South Korea or those transported from other countries would affect the air quality; those data should have been included in variables to reach a solid conclusion. Unfortunately, this issue may be beyond the scope of this study because the data were not accessible from any online site in this study. One thing to point out is that temperature inversions would also affect the air quality in the winter season in Seoul. As cooler air is entrapped below the inversion layer where warmer air is located, air stagnation would be aggravated in the winter, resulting in horrible air pollution in Seoul, especially in the winter.

Table 2: The difference in $PM_{2.5}$ levels between westerly and non-westerly winds in 2019. $PM_{2.5}$ levels were higher in the westerly wind than non-westerly wind in the entire 2019 year, whereas they were not when sub-analyzed in the winter and early spring.

Overall (365 days in 2019)	West		Non-west		P-value
	No.	PM2.5 [μg/m³] [Median (25th, 75th)]	No.	PM2.5 [μg/m³] [Median (25th, 75th)]	
Dominant wind direction	161	23.4 (16.9, 35.5)	204	19.4 (12.0, 26.2)	< 0.001
Prevailing wind direction	180	23.4 (17.5, 34.1)	185	18.7 (11.2, 25.7)	< 0.001
Winter and Early Spring (114 days in 2019)	**West**		**Non-west**		**P-value**
	No.	PM2.5 [μg/m³] [Median (25th, 75th)]	No.	PM2.5 [μg/m³] [Median (25th, 75th)]	
Dominant wind direction	51	32.9 (23.0, 53.5)	39	31.8 (22.2, 39.5)	0.390
Prevailing wind direction	63	23.4 (22.9, 48.0)	27	34.4 (21.2, 47.3)	0.846

Meteorological Factors Associated with Poor Air Quality:

Because the prevailing wind direction in South Korea is westerly, particles originating from neighboring countries to the west, such as China, could influence ambient air pollution levels.[10] Furthermore, the relatively high $PM_{2.5}$ and PM_{10} levels observed in winter (Figure 1) could be attributed to increased air pollution due to burning fossil fuels for heating.[14,23]

The statistical difference between meteorological variables associated with $PM_{2.5}$> 25 μg/m³ and $PM_{2.5}$≤ 25 μg/m³ was investigated through univariate analysis using the 2019 dataset (Table 3). The Wilcoxon rank-sum test was used for precipitation duration and amount, maximum and average wind speed, average relative humidity, vapor pressure, and temperature. Fisher's exact test was used to determine the dominant and prevailing wind directions. The test resulted in P-values< 0.1 for the precipitation amount, maximum wind speed, average wind speed, average vapor pressure, average temperature, dominant wind direction, and prevailing wind direction. Using variables with P-values< 0.1, a multiple logistic regression was subsequently conducted to identify factors linked to elevated $PM_{2.5}$ levels. This resulted in P-values< 0.05 for average wind speed, average vapor pressure, and prevailing wind direction. Lower wind speed, lower vapor pressure, and westerly prevailing wind direction were significantly associated with elevated $PM_{2.5}$ levels.

Table 3: Univariate and multivariate analysis of meteorological factors affecting air quality in 2019. Average wind speed, average vapor pressure, and prevailing wind direction were independent factors associated with poor air quality in the entire 2019 year.

	$PM_{2.5}$ ≤25 μg/m³ (n=238)	$PM_{2.5}$ >25 μg/m³ (n=127)	P-value	Multiple logistic regression		
				P-value	Odds ratio	95% CI
Precip. duration (hr)	0.0 (0.0, 2.8)	0.0 (0, 1.25)	0.156[1]			
	2.5 ± 4.7*	1.5 ± 3.5*				
Precipitation (mm)	0.0 (0.0, 0.6)	0.0 (0.0, 0.0)	0.053[1]			
	3.3 ± 9.8*	0.8 ± 2.6*				
Max. wind speed (m/s)	4.3 (3.6, 4.9)	4.1 (3.5, 4.7)	0.062[1]			
Avg. wind speed (m/s)	1.9 (1.6, 2.4)	1.7 (1.4, 2.1)	< 0.001[1]	**<0.001**	0.33	(0.20, 0.53)
Avg. relative humidity (%)	57.5 (45, 67)	56.1 (46, 66)	0.618[1]			
Avg. vapor press. (hPa)	10.9 (4.7, 19)	5.9 (4.0, 11)	< 0.001[1]	**<0.001**	0.92	(0.89, 0.96)
Avg. temperature (°C)	17.5 (7.3, 24)	7.8 (2.7, 20)	< 0.001[1]			
Dominant wind direction			0.006[2]			
Non-west	146 (61.3%)	58 (45.7%)				
West	92 (38.7%)	69 (54.3%)				
Prevailing wind direction			0.002[2]	**0.017**	1.80	(1.11, 2.92)
Non-west	135 (56.7%)	50 (39.4%)				
West	103 (43.3%)	77(60.6%)				

Note-Data are presented as median (interquartile range) for continuous variables and number (percentage) for categorical variables. *indicates mean ± standard deviations. [1] P-value based on Wilcoxon rank-sum test;[2] P-value based on Fisher's exact test

To identify the meteorological factors affecting $PM_{2.5}$levels, it is desirable to analyze data from one season or similar climatic conditions. Thus, an identical analysis to that described above was conducted using data from the January–March period (Table 4). The Wilcoxon rank-sum test and Fisher's exact test statistics showed that the P-values for precipitation, maximum wind speed, average wind speed, relative humidity, vapor pressure, and temperature were less than 0.1. A multiple logistic regression was subsequently conducted using variables with a P-value< 0.1. This resulted in P-values< 0.05 for maximum wind speed and vapor pressure.

Table 4: Univariate and multivariate analysis of meteorological factors affecting air quality between January and March 2019. Maximum wind speed and vapor pressure were independent factors associated with poor air quality.

	$PM_{2.5}$ ≤25 μg/m³ (n=32)	$PM_{2.5}$ >25 μg/m³ (n=58)	P-value	Multiple logistic regression		
				P-value	Odds ratio	95% CI
Precip. duration (hr)	0.0 (0.0, 0.6)	0.0 (0.0, 0.5)	0.996[1]			
	1.0 ± 2.7*	1.5 ± 3.8*				
Precipitation (mm)	0.0 (0.0, 0.0)	0.0 (0.0, 0.0)	0.477[1]			
	0.1 ± 0.4*	0.8 ± 3.1*				
Max. wind speed (m/s)	4.7 (3.6, 5.8)	4.2 (3.2, 5.2)	0.064[1]	**0.015**	0.49	(0.28, 0.87)
Avg. wind speed (m/s)	2.1 (1.7, 2.5)	1.7 (1.3, 2.1)	0.009[1]			
Avg. relative humidity (%)	39.6 (32, 46)	51.6 (41, 60)	< 0.001[1]			
Avg. vapor press. (hPa)	2.1 (1.6, 3.3)	4.1 (3.2, 5.2)	< 0.001[1]	**<0.001**	2.81	(1.74, 4.55)
Avg. temperature (°C)	-1.5(-3.5, 3.3)	4.0 (1.3, 7.1)	< 0.001[1]			
Dominant wind direction						
Non-west	13 (40.6%)	26 (44.8%)	0.871[2]			
West	19 (59.4%)	32 (55.2%)				
Prevailing wind direction						
Non-west	8 (25.0%)	19 (32.8%)	0.597[2]			
West	24 (75.0%)	39 (67.2%)				

Note-Data are presented as median (interquartile range) for continuous variables and number (percentage) for categorical variables. *indicates mean ± standard deviations. [1] P-value based on Wilcoxon rank-sum test;[2] P-value based on Fisher's exact test

The results of the above test indicated that maximum wind speed was one of the crucial factors affecting the elevated $PM_{2.5}$ levels. The maximum wind speed was slower on days with relatively high $PM_{2.5}$ levels. This result is closely related to that of a previous study, in which severe air pollution episodes were often observed alongside stagnant air conditions.[24] This phenomenon is exacerbated in winter when the low planetary boundary layer inhibits air pollution dispersion.[1,25] The Wilcoxon rank-sum test showed that temperature is also related to PM levels (P< 0.001). High-pressure systems in Siberia often result in cold surges, which rapidly decrease temperature and blow PM away. Elevated $PM_{2.5}$ levels are also correlated with high vapor pressure, which tends to increase with temperature. However, despite this correlation, the temperature was not a statistically significant variable in the multiple logistic regression (Table 4).

Nevertheless, this research has several limitations. Firstly, this study focused on meteorological data to evaluate factors linked with $PM_{2.5}$ without considering emissions from South Korea or those transported from other countries. Secondly, the link between $PM_{2.5}$ levels and meteorological variables could not be supported scientifically with statistical significance. For example, although wind direction was not among the statistically significant variables, it should have been included, given the significant emissions originating west of South Korea. To overcome these limitations, emissions of air pollutants and meteorological variables in both South Korea and surrounding countries should be analyzed together in the future. Trajectory or factor analysis (e.g., positive matrix factorization, chemical mass balance, etc.) with a speciated particulate concentration should also be used to improve the accuracy of the analysis. In addition, data analysis, including many cumulative years, is needed to gather sufficient samples for future studies from January to March.

■ Conclusion

The relationship between meteorological variables (precipitation duration and amount, maximum and average wind speed, dominant and prevailing wind direction, relative humidity, vapor pressure, and temperature) and elevated $PM_{2.5}$ levels in Seoul were evaluated using an annual dataset for 2019 and a dataset for January, February, and March when $PM_{2.5}$ levels were relatively high. The analysis of the 2019

dataset revealed that the most common dominant wind direction was westerly and was associated with elevated $PM_{2.5}$ levels. In contrast, the three-month sub-analysis did not show any statistically significant differences in median $PM_{2.5}$ levels between westerly and non-westerly winds. The multiple logistic regression conducted during the January–March dataset analysis demonstrated that maximum wind speed and vapor pressure were significant independent predictors of elevated $PM_{2.5}$ levels.

Acknowledgments

I would like to thank professor Soohyun Ahn for her assistance with the statistical analysis.

References

1. Kim, H.C.; Kim, S.; Kim, B.-U.; Jin, C.-S.; Hong, S.; Park, R.; Son, S.-W.; Bae, C.; Bae, M.; Song, C.-K. Recent increase of surface particulate matter concentrations in the Seoul Metropolitan Area, Korea. *Sci Rep* **2017**, 7, 1-7
2. Han, C.; Kim, S.; Lim, Y.-H.; Bae, H.-J.; Hong, Y.-C. Spatial and temporal trends of number of deaths attributable to ambient PM2.5 in the Korea. *J Korean Med Sci* **2018**, 33.
3. Choi, J.; Oh, J.Y.; Lee, Y.S.; Min, K.H.; Hur, G.Y.; Lee, S.Y.; Kang, K.H.; Shim, J.J. Harmful impact of air pollution on severe acute exacerbation of chronic obstructive pulmonary disease: particulate matter is hazardous. *INT J Chronic Obstr* **2018**, 13, 1053.
4. Kim, K.-H.; Shon, Z.-H. Long-term changes in PM10 levels in urban air in relation with air quality control efforts. *Atmos environ* **2011**, 45, 3309-3317.
5. Han, S.H.; Kim, Y.P. Long-term Trends of the Concentrations of Mass and Chemical Composition in PM 2.5 over Seoul. *J Korean Soc Atmos Environ* **2015**, 31, 143-156.
6. Kim, Y.P.; Lee, G. Trend of air quality in Seoul: Policy and science. *Aerosol Air Qual Res* **2018**, 18, 2141-2156.
7. Ryou, H.; Heo, J.; Kim, S.-Y. Source apportionment of PM10 and PM2.5 air pollution, and possible impacts of study characteristics in South Korea. *Environ pollut* **2018**, 240, 963-972.
8. Yang, Y.; Liao, H.; Lou, S. Increase in winter haze over eastern China in recent decades: Roles of variations in meteorological parameters and anthropogenic emissions. *J Geophys Res Atmos* **2016**, 121, 13050-13065.
9. Oh, H.-R.; Ho, C.-H.; Koo, Y.-S.; Baek, K.-G.; Yun, H.-Y.; Hur, S.-K.; Choi, D.-R.; Jhun, J.-G.; Shim, J.-S. Impact of Chinese air pollutants on a record-breaking PMs episode in the Republic of Korea for 11–15 January 2019. *Atmos Environ* **2020**, 223, 117262.
10. Park, H.; Lim, W.; Oh, H. Cross-Border Spillover Effect of Particulate Matter Pollution between China and Korea. *Korean Econ Rev* **2020**, 36, 227-248.
11. Shin, M.; Kang, Y.; Park, S.; Im, J.; Yoo, C.; Quackenbush, L.J. Estimating ground-level particulate matter concentrations using satellite-based data: a review. *GIScience & Remote Sens* **2020**, 57, 174-189.
12. Park, S.; Shin, M.; Im, J.; Chang-Keun, S.; Choi, M.; Kim, J.; Lee, S.; Park, R.; Kim, J.; Dong-Won, L. Estimation of ground-level particulate matter concentrations through the synergistic use of satellite observations and process-based models over South Korea. *Atmospheric Chem Phys* **2019**, 19, 1097-1113.
13. Kim, M.J. Changes in the Relationship between Particulate Matter and Surface Temperature in Seoul from 2002–2017. *Atmosphere* **2019**, 10, 238.
14. Ahmed, E.; Kim, K.-H.; Shon, Z.-H.; Song, S.-K. Long-term trend of airborne particulate matter in Seoul, Korea from 2004 to 2013. *Atmos Environ* **2015**, 101, 125-133.
15. KMA, Meteorological Data Portal. http://data.kma.go.kr 2021
16. AirKorea, Air Quality Data. https://www.airkorea.or.kr/eng 2021
17. WHO. Air quality guidelines: global update 2005: particulate matter, ozone, nitrogen dioxide, and sulfur dioxide; *World Health Organization*: **2006**.
18. Astel, A; Tskovski, S; Barbieri, P; Simeonov, V. Comparison of self-organizing maps classification approach with cluster and principal components analysis for large environmental data sets. *Water Res* **2007**, 41, 4566-4578.
19. Becker, S; Soukup, J; Sioutas, C; Cassee, F. Response of human alveolar macrophages to ultrafine, fine, and coarse urban air pollution particles. *Experi Lung Res* **2003**, 29, 29-44.
20. Skov, T; Cordtz, T; Jensen, L; Saugman, P; Schmidt, K; Theilade, P. Modifications of health behavior in response to air pollution notifications in Copenhagen. *Soc Sci Med* **1991**, 33, 621-626.
21. Zhu, Z; Fan, Z; Wu, X; Meng, Q; Wang, S; Tang, X; Ohman-Strickland, P; Georgopoulos, P; Zhang, J; Bonanno, L; Held, J; Lioy, P. Spatial variation of volatile organic compounds in a "Hot Spot" for air pollution. *Atmos Environ* **2008**, 42, 7329-7338.
22. Korn, E; Whittemore, A. Methods for analyzing panel studies of acute health effects of air pollution. *Biometrics*. **1979**, 35, 795-802.
23. Lee, S.; Ho, C.-H.; Choi, Y.-S. High-PM10 concentration episodes in Seoul, Korea: Background sources and related meteorological conditions. *Atmos Environ* **2011**, 45, 7240-7247.
24. Baek, K.-M.; Seo, Y.-K.; Kim, J.-Y.; Baek, S.-O.; Baek, K.-M.; Seo, Y.-K.; Kim, J.-Y.; Baek, S.-O. Monitoring of particulate hazardous air pollutants and affecting factors in the largest industrial area in South Korea: The Sihwa-Banwol complex. *Environ Eng Res* **2019**, 25, 908-923.
25. Chan, K.M.; Wood, R. The seasonal cycle of planetary boundary layer depth determined using COSMIC radio occultation data. *J Geophys Res Atmos* **2013**, 118, 12422-12434.

Author

Yubeen Lee is a senior at Saint Paul Preparatory Seoul High School, Korea. She is interested in studying mathematics and all fields of science and hopes to major in statistics. Her experience in adolescent idiopathic scoliosis and surgery led her to a deep interest in human rights and volunteer work. She is actively involved in human rights activities, serving as an Amnesty PLUS editor and National Honor Society member.

DOI: 10.36838/v4i6.21

Marine Biology Research at Bahamas

Unique and exclusive partnership with the Gerace Research Center (GRC) in San Salvador, Bahamas to offer marine biology research opportunities for high school teachers and students.

- Terra has exclusive rights to offer the program to high school teachers and students around world.
- All trips entail extensive snorkeling in Bahamian reefs as well as other scientific and cultural activities.
- Terra will schedule the program with GRC and book the flights from US to the GRC site.
- Fees include travel within the US to Island, lodging, meals, and hotels for transfers, and courses.
- For more information, please visit terraed.org/bahamas.html

Terra is a N.Y. based 501.c.3 non-profit organization
dedicated for improving K-16 education

IJHSR International Journal of High School Research

is a publication of

terra science and education

N.Y. based 501.c.3 non-profit organization
dedicated for improving K-16 education