Pecan: A Novel Approach to Energy
Supply and Demand Forecasting in a
Photovoltaic Microgrid
Page 27

CJSJ

STAFF

Editor-in-Chief
Mayeesa Rahman, BC'23

Review Editors
Teodora Bratu, CC'24
Gloria Charite, CC'23
Elisa Huang, CC'24
Lindsey Linxi Yang, CC' 24
Jimmy Zhang, CC' 25
Andrew Zheng, SEAS'24

Writers
Yaqoub Ahmad
Jessamine Qu
Penelope Strong
Alice Feng
Kunwoo Kim
Lyon Kim
Vaibhav Mishra
Ishika Nag
Ryan Park
Shreya Sreekantham
Eric Wang
Andy Xu
Krithika Karthik
Francesca Froio
Mirika Jambudi

Special thanks to the general body members of the Columbia Undergraduate Science Journal for their help in reviewing submissions!

Table of Contents

CJSJ

Dear Readers,

I am proud to announce the publication of the 2021-2022 edition of the *Columbia Junior Science Journal*! This year's edition features research in a wide variety of scientific disciplines, including biology, medicine, environmental science, astrophysics, engineering, and chemistry. The quality of research and the scholarship of our authors is impressive - including their work in our publication is an honor.

The last year has been quite a roller coaster. The COVID-19 pandemic forced many to work remotely, reducing scientific research opportunities for students around the world. This happened at a time when scientific literacy and research was needed the most. So, it is inspiring to know that so many young scientists across the world are advocating for science through their research and authorship as we continue to rise above the pandemic. In fact, we had a record-high number of submissions to the *Columbia Junior Science Journal* this year!

Science is a collective endeavor. So, in addition to advocating for scientific research, it is important to increase the accessibility of scientific resources to all communities. Giving high school students a platform to contribute to our scientific understanding of nature is only one way we seek to fulfill that aim. Our science events, conferences, and outreach efforts are a few additional ways we try to increase accessibility of scientific involvement to all. I am proud of our editorial team for leading the *Columbia Undergraduate Science Journal* in this cause.

As the President of the *Columbia Undergraduate Science Journal*, it was an honor to read, edit, and review submissions made to our journal. I am grateful to our *Columbia Junior Science Journal* team and editorial board for their significant contributions to our scientific review process. I am especially grateful to Mayeesa Rahman, Editor-in-Chief of the *Columbia Junior Science Journal*, whose leadership made this publication a success. I would also like to thank the *Columbia Undergraduate Science Journal* Faculty Advisory Board, a group of esteemed Columbia University professors whose support ensures publication of the highest quality of scholarship.

Congratulations to our authors, and thank you to our readers!

Arjun Kudinoor
President, Chief Editorial Officer
Columbia Undergraduate Science Journal

CJSJ

Dear Readers,

I am very excited to announce the publication of the seventh edition of the *Columbia Junior Science Journal*! As editors of this journal, our goal has always been to highlight outstanding research completed by high school students, and this issue features a collection of impressive and novel work in a wide range of scientific disciplines.

I would like to take this opportunity to thank everyone who was involved in the creating this edition: the *Columbia Junior Science Journal* editorial board, the student authors featured, their research mentors, their parents and peers, and anyone else who contributed. To the authors – we value your unwavering commitment to scientific discovery and excellence in these challenging times and appreciate your hard work both while completing your research and working with us during the editing process.

During this tumultuous and confusing year as we continue to navigate the COVID-19 pandemic, we recognize how difficult it has been to obtain the resources and laboratory access necessary for hands-on research. Due to these unfortunate circumstances and our dedication to making scientific publication as accessible as possible, we have decided to accept both original research articles and review papers for the second year in a row. We are proud to present both types of research in this issue. We received a record number of over 350 submissions this year, and we are delighted that the expansion of our journal has granted more students the opportunity to submit their research.

This year's edition of the *Columbia Junior Science Journal* explores a broad spectrum of scientific innovation ranging from modeling the kinematics of a galaxy to developing a novel program for the detection of lung carcinoma. We hope that you read this set of articles as one cohesive publication and can be inspired, as we at CJSJ are, by the endless potential of the next generation's incredible scientific thinkers.

It has been a great pleasure serving as the 2021-2022 Editor-in-Chief of CJSJ and watching our editorial board, as well as the high school students, grow throughout this entire publication process. I look forward to seeing and reading about all of your continued scientific endeavors for many years to come!

Sincerely,

Mayeesa Rahman
Editor-in-Chief
Columbia Junior Science Journal

Evaluation of Plastic Materials for Football Helmet Facemasks Using Finite Element Simulation

Kunwoo Kim

Abstract— **American football reports the highest rate of head injuries including concussion in the United States. Although the use of football helmets has protected players and reduced severe head injuries, the number of concussion incidents has not decreased. A facemask mounted on the front opening of the helmet to protect the face is the second most impacted location of concussions. However, its high-rigidity material such as carbon steel or titanium makes the facemask disadvantageous in absorbing impact. This study aimed to assess the feasibility of using a plastic facemask to absorb impact in order to prevent concussion. Four different plastics, Polycarbonate (PC), Polyetherimide (PEI), Polymethyl Methacrylate (PMMA), and Acrylonitrile Butadiene Styrene (ABS), were selected to make the facemask, and a pneumatic ram impact test was conducted using numerical simulation. A PEI facemask was deformed but withstood the ram impact while preventing the ram from hitting the face. On the other hand, ABS and PMMA facemasks were cracked and failed, and a PC facemask was not cracked but allowed the ram to hit the face. This study indicated that if the plastic material was optimized, a PEI facemask would be feasible for absorbing the severe external impact, and protect players from concussions.**

I. INTRODUCTION

American football is the most popular sport in the United States, but it has the highest head injury rate among all sports. Despite the National Football League(NFL)'s tremendous effort, the number of concussion incidents has not decreased since 2015 [1]. Football players wear helmets that consist of an outer shell, paddings, clips, a chinstrap, and a facemask (Figure 1). The structure and material of the outer shell and padding have been studied to improve impact-absorbing performance for a long time: plastic has been applied to the outer shell. The facemask which is for face protection, however, is the second most impacted location of concussion following the side of the helmet according to NFL [2]. little research has been done on facemasks for preventing head injuries. Modern facemasks are made of metal such as carbon steel or titanium. Although a metal facemask is stiff and effective to protect the face from direct contact damage, it can transfer the impact force directly to the head and increase the risk of concussion. The purpose of this study was to evaluate the feasibility of applying a plastic material to the facemask for absorbing impact and preventing concussion by investigating various plastic materials with finite element simulation.

.

II. METHOD

The NFL requires the test condition of the National Operation Committee on Standards for Athletic Equipment (NOCSAE) for the helmet. One of the test conditions is the pneumatic ram impact test [3,10]. The ram consists of a front plastic cap, foam, and steel rod, and weighs 15.6 kg (Figure 1). Finite element simulation, a method commonly used for impact simulation, was conducted based on this test condition. The ram was set to impact to the front of the facemask at a speed of 7.4 m/s, the threshold impact speed for a concussion in NFL [4].

The NFL provides open-source simulation data of head-neck and football helmet models [5]. The 3-dimensional facemask and ram impactor model were recreated in detail (Figure 1) from the open-source model, in which the facemask was realized as a simplified one-dimensional model. The updated facemask had a circular section of 8 mm diameter. Since actual facemasks are attached to the helmet using clips, clip areas of the facemask were fixed as the boundary condition in the simulation (Figure 1).

Four plastic materials, Polycarbonate (PC), Polyetherimide (PEI), Polymethyl methacrylate (PMMA), and Acrylonitrile butadiene styrene (ABS), were selected for the simulation. PC and ABS are currently used for the helmet outer shell due to their excellent ductility. PC has the best ductility with 110 % elongation at break. PEI is famous for ultra-performance engineering plastic and exhibits the highest strength (110 MPa) with the second-best elongation of 50 %. PMMA, also known as Acrylic, is a common plastic material has the best stiffness equivalent to PEI, and the second-highest strength (67 Mpa). The stress-strain curves of each material were acquired from CAMPUS Plastics (Figure 2) [6,7,8,9]. These stress-strain curves were entered into the simulation to realize nonlinear material behavior.
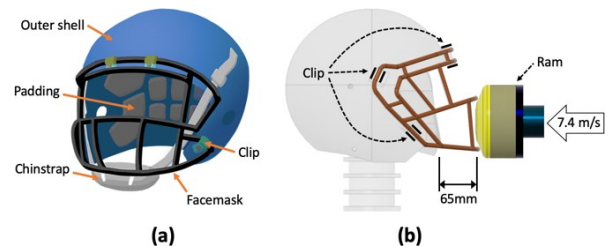

Figure 1. (a) Football helmet composition (b) Ram impact test condition and simulation model (Facemask and ram)
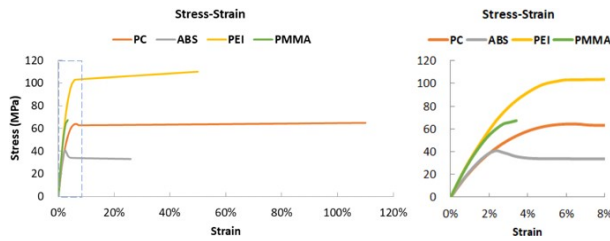
Figure 2. Stress-stain curve (Right curve: Strain from 0 to 8%) - Represent the material behavior against external force

Impact simulation was conducted with an impact event time of up to 0.025 seconds when the ram almost stopped. The contact condition between the facemask and the impactor was defined with zero friction. In the finite element simulation setting, a second-order element option was used for higher accuracy, and elements that reached the maximum tensile strain of the material during the event and were not able to resist against force set to be deactivated to realize crack propagation.

III. RESULTS AND DISCUSSION

It was required that the facemask was not broken by the ram impact and it prevented the ram from hitting the face during the event of the impact. So, the facemasks were examined for cracks until the ram slowed down sufficiently, and the ram velocity and displacement over time were measured from the simulation (Figure 4).

Figure 3 shows the deformation of the PEI facemask for up to 0.025 seconds when the ram was almost stopped (Figure 3,4). The facemask did not break and prevented the ram from touching the face. On the other hand, ABS and PMMA facemasks cracked at 0.014 and 0.002 seconds and failed to protect the face (Figure 5). The ABS facemask could not fully absorb impact energy even with a good ductility of 26 % elongation due to the lowest material strength. The PMMA facemask showed the worst result due to its brittleness of 3.4 % elongation despite its excellent stiffness. The PC facemask was not cracked like PEI facemask, but was deformed by more than 65 mm, allowing the ram to hit the face slightly due to its lower stiffness than PEI (Figure 4.5). These results in failure mode or deformation indicate that the performance of the facemask to absorb impact and protect the face was highly dependent on the material properties.
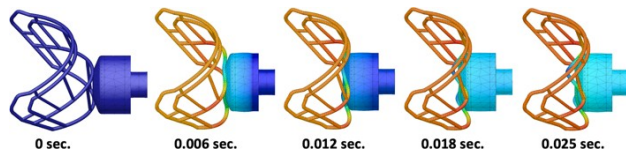

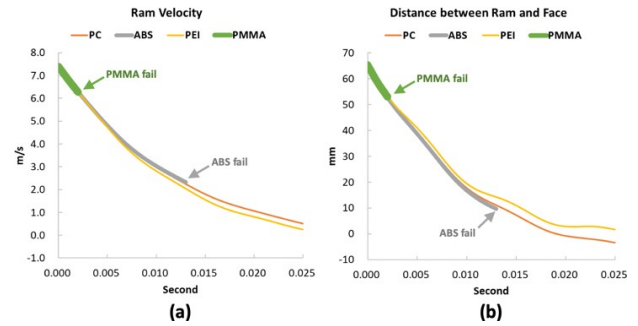Figure 3. Deformation of PEI facemask over time

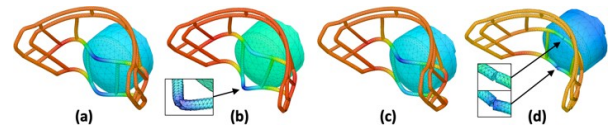Figure 4. (a) Ram velocity (b) Distance between ram and face

Figure 5. (a) max deformation of PC facemask (b) failure of ABS facemask (c) max deformation of PEI facemask (d) failure of PMMA facemask

IV. CONCLUSION

The facemask on a football helmet is the second most impacted position causing concussions in professional football, but the current facemasks have a metal frame structure that is too stiff to absorb the impact and gives direct impact to the head. Therefore, extensive simulations were conducted to evaluate plastic materials for the facemask to mitigate the risk of concussion. Plastic is flexible and absorbs impact better than metal, so four different plastic materials, PC, ABS, PEI, and PMMA, were validated using simulation based on the ram impact test condition of 7.4 m/s ram impact velocity, the concussion threshold speed. ABS and PMMA facemasks cracked and failed, while a PC facemask didn't fail but was too deformed to protect the face. A PEI facemask was deformed as it absorbed the impact, but it succeeded in protecting the face. This study suggested that a PEI facemask endured a harsh ram impact condition while absorbing the impact more than a metal facemask, and this indicates that a plastic facemask is feasible for an actual helmet to mitigate the concussion risk if the plastic material and frame structure design are further optimized for different ram attack angles and speeds

V. REFERENCES

[1] G. O National Football League (NFL). (2021). Injury data since 2015. https://www.nfl.com/playerhealthandsafety/health-and-wellness/injury-data/injury-data

[2] National Football League (NFL). (2018). Video review webinar. https://www.nfl.com/photos/video-review-webinar#9921dfe5-489e-4632-867e-23a6e24d2b37

[3] National Operation Committee on Standards for Athletic Equipment (NOCSAE). (2019). Standards Performance Specification For Newly Manufacture Football Helmets, NOCSAE Doc (ND)002-17m19.

[4] Pellman, E., Viano, D., Withnall, C., Shewchenko, N., Bir, C., Halstead, P. (2006). Concussion In Professional Football: Helmet Testing To Assess Impact Performance - Part II. Neurosurgery, 58, 78–96.

[5] National Football League (NFL). (2018). Finite Element Models: New Tools for Innovation in Football Helmet Design, from

https://www.nfl.com/playerhealthandsafety/equipment-and-innovation/engineering-technology/finite-element-models-new-tools-for-innovation-in-football-helmet-design

[6] CAMPUS®. 2020. Makrolon® 2405, from https://www.campusplastics.com/campus/en/datasheet/Makrolon%C2%AE+2405/Covestro/22/986598f8

[7] Dundar, M. (2017). Strain Rate Dependence And Impact Behavior Of Abs (acrylonitrile-Butadiene-Styrene) Amorphous Thermoplastic [Doctoral dissertation, Wayne State University].

[8] CAMPUS®. 2020. PLEXIGLAS® 6N, from https://www.campusplastics.com/material/pdf/133627/PLEXIGLAS6N?sLg=en

[9] GE Plastic. 2003. ULTEM PEI Resin Product Guide.

[10] Gunnarsdóttir A. (2019). Evaluation of Test Methods for Football Helmets Using Finite Element Simulations [Degree Project in Medical Engineering, KTH Royal Institute of Technology]

A Deep Learning Model for Protein Abundance Prediction from RNA Data with Manifold-Preserving Regularization

Alice Feng, Shaoheng Liang, Ken Chen Ph.D.

Abstract—**A key challenge in single-cell multiomics study is to quantify the relationship between mRNA level and protein abundance. This relationship is complicated by the dynamic nature of mRNA and protein. In this paper, a deep learning regression model was proposed to predict protein abundance from mRNA expression data. However, overfitting was identified as a major source of error. Because different modalities of the same sample should concord to the same cell population structure, we invented a manifold-preserving regularization term to reduce overfitting induced by noise specific to training. By applying our model on CITE-seq data for 25 cell-surface proteins representing well-characterized markers, we observed an improvement of up to 30% on testing error. Thus, manifold-preserving regularization helps distill true mRNA-protein relationships from noisy data. We expect it to be generally applicable to other multiomics applications.**

I. Introduction

Single-cell multiomics technologies provide co-assayed measurements from the same cells and can help learn relationships among different omics. Protein and RNA are important in understanding how cells work, and studying their relationship provides further insight into disease progression. A key challenge in single-cell multiomics study is to quantify the relationship between mRNA level and protein abundance [1,2], which is complicated by the dynamics in synthesis, splicing and degradation of mRNAs, and modification, folding and transportation of proteins. Furthermore, co-assayed data are limited, as they are usually more costly and compromise throughput and read depth. Hence, it is desirable to find a computational method that quantifies the relationship between proteins and RNA and predicts the protein profile with certain accuracy.

Overparameterized deep learning models are ideal for modeling complicated relationships, including that of mRNAs and proteins. However, overfitting occurs when a model fits to not only the true relationship of the predictors and responses, but also training data-specific noise (such as technical noise from sample preparation and sequencing). To reduce the sensitivity to noise, regularization is a common way to shrink the solution space using heuristics.

For single-cell data, manifold is an important characteristic that represents the cell population structure in a biospecimen. The manifold of a single-cell dataset is defined by the pairwise similarities of cells that characterize clusters, subtypes, and trajectories. It has been widely used in dimension reduction and trajectory inference [3,4,5]. Because different modalities

* Alice is with the Harker School, San Jose, CA. Shaoheng Liang and Dr. Ken Chen are with Depart of Bioinformatics and Computational Biology at the University of Texas MD Anderson Cancer Center, Houston. TX.

of the same sample largely agree, we hypothesize that the majority of discrepancies between the manifolds of mRNA and protein data are technical noise. To mitigate overfitting, we introduce Manifold-Preserving Regularization (MPR), which suppresses the noise to mitigate overfitting (Fig. 1).

In this paper, a deep learning regression model was built to map the relationship between protein and gene expression levels. For a specific protein, although its coding gene can be a predictor, other genes also play important roles in the complicated regulations of its production and degradation. Thus, we use all genes as predictors. Results show that MPR increases the accuracy of models, compared to unregularized L2-regularized models. The genes with high weights in the resulting model are indeed in related biological pathways.
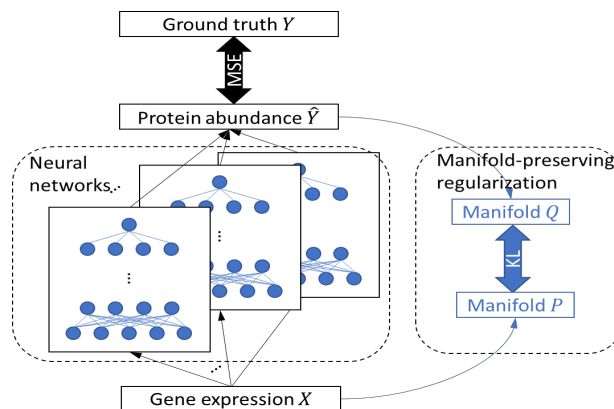
Figure 1. Architecture of Deep Neural Network with MPR

II. Methods

We used a CITE-seq dataset with co-assayed bone marrow mononuclear cells from two healthy donors [6]. We used 14,468 cells from Donor 1 as the training data, and, to test if the model can generalize, 16,204 cells from Donor 2 as the testing data. Measured in the datasets are 25 proteins and 17,009 genes. Twenty-seven cell types were identified by the original publication.

First, we preprocessed the data, including data cleaning and normalization. For the protein data, we normalized, log-transformed, and scaled the data. For the RNA data, we first normalized and performed log transform using the same process as with the protein data. Based on our analysis of the

data, we decided the most effective data cleaning method was removing genes with constant or similar values across cell samples. As a common practice in single-cell data analysis, we used a cutoff of dispersion at 0.5 and mean at 0.5, 1,555 highly variable genes are selected for Donor 1 (and later also used on Donor 2). Finally, we also scaled the RNA data.

Then the goal of our regression analysis is to identify a function of the RNA expression data $\mathbf{X}$ so that the function value is close to the protein abundance value $\mathbf{Y}$, let the function be $f(\mathbf{X},\beta)$ where β is the function parameters, we need to optimize the function form and the parameters so that the prediction $\hat{\mathbf{Y}}=f(\mathbf{X},\beta)$ is as close to $\mathbf{Y}$ as possible.

We selected a 5-layer neural network with 1500 nodes per hidden layer and a batch size of 64 after experimenting with the number of layers, number of nodes in the first layer, and the batch size. Each protein has its own neural network. Mean squared error (MSE) is used as the loss function. We also measured the Overfitting Factor of each model defined as $1 - \sqrt{\text{Training MSE} \mid \text{Testing MSE}}$.

Based on our hypothesis that matching the manifold of RNA and protein will mitigate overfitting, we used manifold information as a regularization in fitting the model (Fig.1). The manifolds for $\mathbf{X}$ and $\mathbf{Y}$ are denoted $\mathbf{P}$ and $\mathbf{Q}$, defined as

$$p_{ij} = \frac{\exp\left(-\left(\left|\left|x_i - x_j\right|\right| - \rho_i\right)/\sigma_i\right)}{\sum_{k \neq l} \exp\left(-\left(\left|\left|x_k - x_l\right|\right| - \rho_i\right)/\sigma_i\right)}$$

and

$$q_{ij} = \frac{\left(1 + \left(\left|\left|\hat{y}_i - \hat{y}_j\right|\right| - \tau_i\right)/\sigma_i\right)^{-1}}{\sum_{k \neq l}\left(1 + \left(\left|\left|\hat{y}_k - \hat{y}_l\right|\right| - \tau_i\right)/\sigma_i\right)^{-1}}$$

where $\rho_i = m \in \left|\left|x_i - x_j\right|\right|, \tau_i = m \in \left|\left|\hat{y}_i - \hat{y}_j\right|\right|$, and σ_i is selected to ensure

$$\sum_j \exp\left(-\left(\left|\left|x_i - x_j\right|\right| - \rho_i\right)/\sigma_i\right) = k$$

The regularization term is defined as the Kullback-Liebler (KL) divergence of P and Q [7], also called relative entropy, measuring the difference between two distributions.

III. Results and Discussion

First, we experimented with deep learning architectures as shown in Tab. 1. Drop-out layers are added to reduce overfitting. Generally, the results are not sensitive to the network architectures. Architectures trained with larger batch sizes tend to have smaller Overfitting Factors. The model with the lowest MSE was model #6: 1500 nodes in the first layer and trained with batch size 64, which is used throughout the rest of this paper.

TABLE I.   MSE COMPARISON AMONG DIFFERENT ARCHITECTURES

Model #	Nodes	Batch Size	MSE	Overfitting Factor
1	500	16	0.372	0.415
2	500	64	0.356	0.279
3	1000	16	0.362	0.472
4	1000	64	0.358	0.314
5	1500	16	0.365	0.505
6	1500	64	0.356	0.344

We trained and tested deep learning models without any regularization on each of the 25 proteins individually. The Overfitting Factor is nearly 1 for all proteins, suggesting that a near-perfect fit was achieved on training data, but does not generalize to testing data. Overfitting is indeed a major problem which necessitates regularization.

We then experimented with L2-regularization and MPR. Both improved the Overfitting Factor significantly. MPR reduced the Overfitting Factor by 25% on average and up to 50% for certain proteins, and outperformed L2 for every protein (Fig 2).
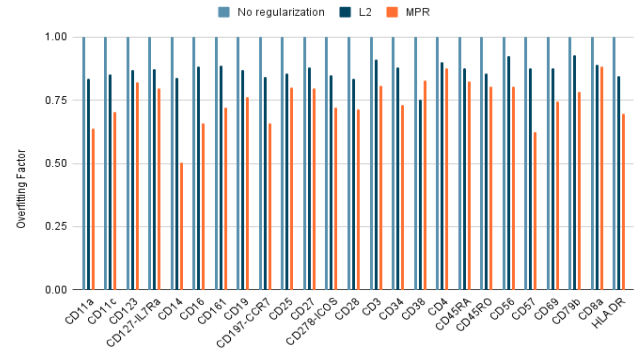
Figure 2.   Overfitting Factor Comparison among No Regularization, L2, MPR

For MSE, MPR consistently reduced the model's testing MSE by 15% on average and up to 30% for certain proteins (Fig. 3). MPR performed preferably for 16 of the 25 proteins compared to L2-regularization, especially for proteins harder to predict from gene expression data, including CD123, CD25, CD38, CD4, CD56, CD79B, and CD8A. When compared with state-of-the-art model cTP-net[2], MRP outperformed cTP-net by up to 15%.

V. REFERENCES

[1] Lill, Christina M., Johannes T. Roehr, Matthew B. McQueen, Fotini K. Kavvoura, Sachin Bagade, Brit-Maren M. Schjeide, Leif M. Schjeide, et al. "Comprehensive Research Synopsis and Systematic Meta-Analyses in Parkinson's Disease Genetics: The PDGene Database."

[2] Meireles, Joana, and João Massano. "Cognitive Impairment and Dementia in Parkinson's Disease: Clinical Features, Diagnosis, and Management."

[3] Nissar, Iqra, et al. "Machine Learning Approaches for Detection and Diagnosis of Parkinson's Disease - A Review."

[4] Solla, Paolo, Antonino Cannas, Federica Carla Ibba, Federico Loi, Marta Corona, Gianni Orofino, Maria Giovanna Marrosu, and Francesco Marrosu. "Gender Differences in Motor and Non-Motor Symptoms among Sardinian Patients with Parkinson's Disease."

[5]Baldereschi, M., A. Di Carlo, W. A. Rocca, P. Vanni, S. Maggi, E.

[6]Perissinotto, F. Grigoletto, L. Amaducci, and D. Inzitari. "Parkinson's Disease and Parkinsonism in a Longitudinal Study: Two-Fold Higher Incidence in Men."

5.

Evaluation of Gender Effects in Predicting Parkinson's Disease from Voice: A Random Forest Approach

Shreya Sreekantham

Abstract— **Parkinson's Disease (PD) is the second most prevalent neurodegenerative disease in the world, affecting more than 10 million people. It has no cure, but early diagnosis of PD can help slow down the progression and improve the patient's life quality. However, the diagnosis of PD is often subjective and inaccurate because its presentation varies widely between individuals. This study focuses on early PD diagnosis and evaluates biomedical voice parameters variation by gender using a novel Random Forest Algorithm (RFA). The study utilized a multivariate PD dataset extracted from the UCI Machine Learning data repository that consisted of 5,875 voice recordings from 42 subjects. The novel RFA introduced in this study both improves the accuracy for PD detection and consistently performs well across gender. In addition, this study identifies gender-based differences in the expression profiles of voice parameters that can be useful in future clinical applications.**

I. INTRODUCTION

Parkinson's Disease (PD) is a condition characterized by a degeneration of dopaminergic neurons in the substantia nigra, impacting motor control, sensory systems, and cognition. [1] Some specific motor symptoms include dysarthria (weakness and coordination issues), akinesia (impairment of voluntary movement), tremors, and vocal changes. PD has no cure, so early diagnosis is critical to monitor, minimize, and control disease progression. However, as symptoms and progression vary widely, PD often goes undiagnosed for many years. In addition, factors like gender differences, aging, and immune status tend to complicate PD diagnosis. Since no diagnostic lab tests for PD exist, traditional diagnosis for PD relies on costly in-clinic tests and are not very accurate to date. For example, a recent study finds that early diagnosis of PD, with the patient having symptoms for five years or less, is only 58% accurate. [2] Therefore, there is an urgent medical need to develop more sensitive tools for PD diagnosis.

Given this subjectivity in diagnosing PD, machine learning approaches that target specific symptoms below the range of human observation have gained popularity. [3] Since vocal impairments are one of the earliest symptoms of the disease, this study focuses on detecting PD from voice. Additionally, despite different presentations in symptoms, there are relatively few studies investigating how gender differences affect the accuracy of PD diagnosis. One significant gender-based difference in Parkinson's disease is that although women demonstrate higher mortality and faster progression rates, PD is twice as prevalent in men than women. [4] This study seeks to understand how gender differences in symptoms affect the voice-based diagnosis of Parkinson's

and proposes a novel Random-Forest Algorithm (RFA) that consistently performs well in detecting PD across gender. Below are the accuracies of PD diagnosis achieved by the algorithm:

- 99.2% for females without gender-impacted parameters
- 96.6% for males without gender-impacted parameters
- 99.3% for females with gender-impacted parameters
- 97.15% for males with gender-impacted parameters

In comparison, previous machine learning approaches for voice-based PD diagnosis when accounting for differences across these genders achieved a highest accuracy of 82.14%. This novel RFA improves PD diagnostic accuracy and performs well across gender. This study is an important step towards creating robust diagnostic models that account for demographic variation.

II. METHODS

Dataset Description

The dataset utilized in this study was generated by Athanasios Tsanas and Max Little of the University of Oxford12. This dataset has 16 biomedical early-stage PD voice parameters from 42 subjects. [5] The data has a total of 5,875 voice recordings during a six-month trial of a telemonitoring device. About 200 recordings were collected from each patient. The data was categorized as follows for each subject: Sex (Male: 0 and Female:1), Jitter (Five measures of variation in fundamental frequency), Shimmer (Six measures of variation in amplitude), NHR (Two measures of noise to tonal components ratio), RPDE (nonlinear dynamical complexity measure), DFA (Signal fractal scaling exponent), and PPE (a nonlinear measure of fundamental frequency variation).

Random Forest Algorithm

The goal of the machine learning algorithm is to predict Unified PD Rating Scale (UPDRS) scores, a scale of PD progression, from the parameters extracted from the voice recordings. Conducting initial exploratory analysis revealed no clear linear, logistic, or planar relationships between variables. To bypass these limitations, a random-forest algorithm (RFA) was utilized. With multiple decision trees rather than just one, Random Forests enables the exploration of the dependencies between variables in the multivariate dataset.

The data was sorted by gender for further analysis. This distribution was then utilized to identify gender-specific patterns of one or a combination of parameters using simpler statistical analysis. Following this Random Forest algorithm (Fig 1) was utilized to further evaluate the prediction parameters.
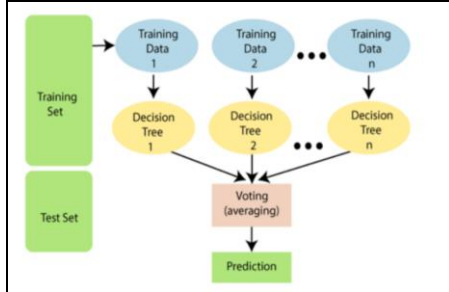

Fig 1. Depiction of Random Forest Cluster

The process for generating the RFA primarily involved the following steps: selecting random data points from the training set, building the decision trees with the selected data points, choosing the number of decision trees, repeating the steps, and finding predictions of each decision tree. Differences based on gender were identified in the data set, and the RFA was applied to the data to compare accuracies in predicting UPDRS scores.

The ensemble form of RFAs is expressed as Forest F = f1, f2, ..., fn. Each F yields a distribution that was averaged prior to applying the algorithm. The predictors were then combined by using the mean of the continuous target variables. The prediction performance of the proposed model was finally confirmed with a 5-fold cross-validation.

III. RESULTS

The dataset utilized in this research article has 5875 recordings, 1867 from females and 4008 from males. Exploratory analysis was conducted using the data from the training instances for all the inputs described in Table 1. Gender based data distributions were compared for both the target variables (motor UPDRS and Total UPDRS). The results showed that both target variables reflected a normal distribution profile (Fig 2).
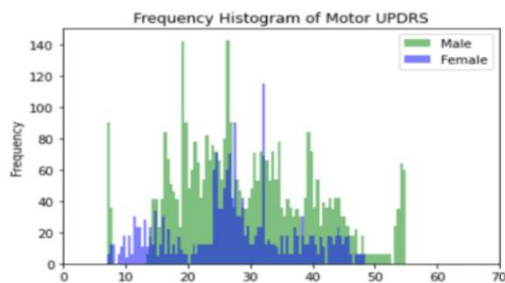

Fig 2. Gender-based frequency histogram profiles for motor UPDRS. Similar patterns were observed for Total UPDRS (not shown)

The biomedical voice parameter measurements were next evaluated individually for distribution trends between male and female subjects. For simplistic analysis, the data was normalized to the male measurements. The analysis established two of the sixteen parameters (Jitter Abs and NHR) showed differences across genders (Fig 3).
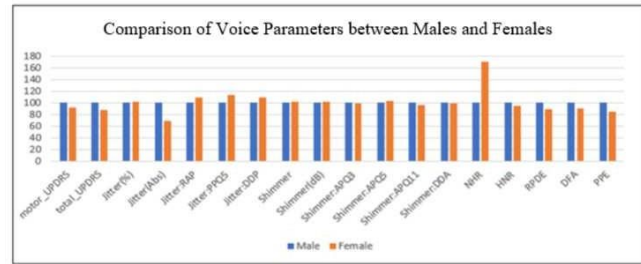

Fig 3. Comparison of individual biomedical parameters establishes possible differences in two of the sixteen parameters. (Data normalized to males)

This data indicates there are gender-based differences in the individual parameters, however it is unclear if these differences affect the prediction of UPDRS scores. Therefore, in order to obtain unbiased results, the RFA was run on the test cohort, to calculate accuracy with and without the gender dependent features (Jitter (Abs) and NHR). These prediction accuracies were then validated with a 5-fold cross-validation. Table 1 below summarizes the accuracies of the RFA.

Table 1: Accuracy Predictions with and without the identified gender-impacted parameters (Jitter Abs and NHS).

Gender	Accuracy Output	W/gender bias parameters	W/o gender bias parameters
Males	Motor UPDRS	97.3	96.7
	Total UPDRS	97.0	96.5
Females	Motor UPDRS	99.4	99.3
	Total UPDRS	99.2	99.1

The RFA averages a 98.0% with or without the gender-impacted parameters. Despite the gender-based differences in the individual parameters, this RFA maintains high accuracy across genders in the dataset.

IV. DISCUSSION

This study introduced a novel RFA which enhanced the accuracy of PD diagnosis compared to existing methods. It also helps conclude that variation in symptoms by gender does not affect the accuracy of voice-based diagnosis of PD. Furthermore, the gender-specific data were evaluated on one generalized model, which reduced the complexity of both the analysis and the prediction. For future work, applying this model to additional datasets and investigating the effects of other demographic differences will be a useful step towards implementing widely applicable voice-based screenings for Parkinson's.

V. ACKNOWLEDGEMENT

VI. REFERENCES

[1] Lill, Christina M., Johannes T. Roehr, Matthew B. McQueen, Fotini K. Kavvoura, Sachin Bagade, Brit-Maren M. Schjeide, Leif M. Schjeide, et al. "Comprehensive Research Synopsis and Systematic Meta-Analyses in Parkinson's Disease Genetics: The PDGene Database."

[2] Beach, Thomas G., and Charles Adler. "Importance of Low Diagnostic Accuracy for Early Parkinson's Disease: Low Diagnostic Accuracy For Early PD."

[3] Solla, Paolo, Antonino Cannas, Federica Carla Ibba, Federico Loi, Marta Corona, Gianni Orofino, Maria Giovanna Marrosu, and Francesco Marrosu. "Gender Differences in Motor and Non-Motor Symptoms among Sardinian Patients with Parkinson's Disease."

[4]Baldereschi, M., A. Di Carlo, W. A. Rocca, P. Vanni, S. Maggi, E. Perissinotto, F. Grigoletto, L. Amaducci, and D. Inzitari. "Parkinson's Disease and Parkinsonism in a Longitudinal Study: Two-Fold Higher Incidence in Men."

[5]"UCI Machine Learning Repository: Parkinsons Telemonitoring Data Set." Accessed December 11, 2021.

Effects of Various Soil Microbiomes on Native and Invasive Plants

Eric Wang

Abstract— **Many nature reserves use exclosures to preserve pockets of native plant biodiversity. These exclosures typically have a high proportion of native plants, as there are few invasive species to harm them, resulting in different plant-soil feedbacks (PSFs) within and outside of exclosures. PSFs alter the soil microbiome and have lasting effects on plant community composition; as such, they have important implications for natural ecosystem conservation. For this case study, three native and two invasive plant species were grown in soil inoculated with microbiomes collected from inside and outside of ten exclosures. Analyses of their biomass revealed that native species performed better in the soil microbiome from the exclosures, while invasive species' growth was not significantly impacted by the different microbiomes. This research provides new insights into plants and the soil microbiome in the context of conservation and has important implications on the protection of natural ecosystems.**

I. INTRODUCTION

The soil microbiome consists of billions of microorganisms living in a complex, underground ecosystem. In relation to their interactions with nearby plants, soil microorganisms can generally be categorized into three groups: enemies, which negatively affect plants; mutualists, which positively affect plants; and decomposers, whose main roles are in nutrient cycles [1]. Taken altogether, the balance of these three groups has significant effects on plant growth and development. Plants can take advantage of these effects by altering the soil microbial community composition to their benefit. For instance, in the "cry for help" phenomenon, plants release various organic chemicals through their roots to recruit certain microorganisms that help them resist environmental stressors, such as drought [2]. These changes are short-term, but plants can also have lasting effects on the soil microbiome. Known as plant-soil feedbacks (PSFs), these changes last long after the original plant is gone and have important influences on plant community assemblage [3].

Plants have species-specific feedback effects, which contribute to differing fitness levels among plant species. In the context of conservation, it is hypothesized that non-native plants may exploit PSFs to outcompete native plants and become invasive. Meisner et al. (2014) recently conducted a meta-analysis of native and non-native PSFs, which revealed that non-native species typically had fewer self-negative PSFs; this may be an important contributor to their invasiveness, which warrants further research into how these PSFs affect the spread of invasive species [5].

Invasive species management is key to preserving native biodiversity and healthy natural ecosystems [6]. To combat the spread of invasive plant species, many nature reserves use exclosures, fenced-in regions meant to protect native biodiversity from herbivores and the spread of invasive species. One study in Europe found that protected areas such as these had lower invasive richness inside compared to outside despite being a suitable habitat for them [7]. A recent survey of the exclosures at the South Mountain Reservation, located in Essex County, NJ, similarly found that exclosures typically had higher native plant coverage than outside the exclosures [8].

This case study analyzes how the soil microbiome inside and outside of the exclosures at the South Mountain Reservation differ in their effects on native and invasive plant species. It seeks to determine whether the PSFs resulting from the differing native plant cover inside and outside exclosures influence how the soil microbiome affects different plant species.

II. METHODS

Soil samples were collected from both inside and outside of ten randomly sampled exclosures in the southeast portion of the South Mountain Reservation. The soil microbiomes of each site were then transferred to a sterilized potting mix of potting soil (80% (v/v)) and topsoil (12.5% (v/v)) through direct inoculation at a rate of 7.5% (v/v), which transfers the soil microbiome while only minimally altering abiotic conditions. [9] A sterile control was included using a sterile inoculant. Pots were incubated overnight at ambient temperature (~27 ℃) prior to planting.

The experiment consisted of 3 blocks and 21 treatments. Two of the blocks had individual plants grown in each pot, while the third block had larger pots with two individuals from each species per pot. Plants were kept in a grow tent under 600 W metal-halide lamps providing approximately 220 μmol light quanta m-2s-1 at plant level with a photoperiod of 16:8 hours (day:night). Ambient temperature was 30℃ day/22℃ night and relative humidity was 50-70%.

Five plant species were chosen based on their family and their relative conservation priorities at the reservation[8]. Three species were native (*Solidago flexicaulis* (Asteraceae), *Calamagrostis canadensis* (Poaceae), and *Geranium maculatum* (Geraniaceae)) and two species were invasive (*Artemisia vulgaris* (Asteraceae) and *Miscanthus sinensis* (Poaceae)). Seeds were surface-sterilized with a solution of 2.75% NaClO and 0.005% Tween20 for 1 minute, then germinated on a moist paper towel medium. Seedlings that died within the first two weeks were replaced, though seedling mortality was low (~1.2%). Plants were watered daily for the first 4 weeks, then every other day for the remainder of the experiment. Plants were allowed to grow for 7 weeks total.

Fresh biomass was then measured as total aboveground biomass; dry biomass was measured by drying the plants.

All statistical analyses were performed using R version 4.1.0 [10]. Biomass data from the smaller pots were analyzed with linear mixed effects models using the *lme4* package with plant species, microbiome origin (inside/outside), and their interaction as fixed effects and a random effect of block. Separate models were also generated for each plant species. Data from the larger pots were analyzed using the total native and total invasive biomasses per pot. Statistical significance was tested using F-tests with the *anova* function in the package *lmerTest*.
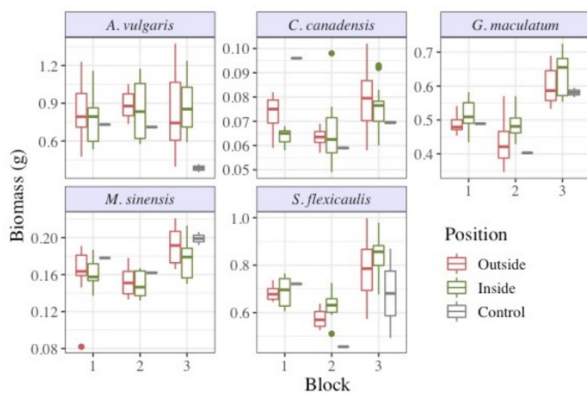
III. RESULTS AND DISCUSSION

Fig. 1 Box plots of biomass separated by species, block, and position

Species	F	p
Overall (Species Position)	$F_{..} = 1.627$	0.1627
A. vulgaris	$F_{..} = 0.3985$	0.5317
C. canadensis	$F_{..} = 2.0298$	0.1626
G. maculatum	$F_{..} = 6.545$	**0.01475**
M. sinensis	$F_{..} = 1.0684$	0.3082
S. flexicaulis	$F_{..} = 3.4441$	0.07146
Total Native Biomass	$F_. = 5.42$	**0.03177**
Total Invasive Biomass	$F_. = 0.2017$	0.6587

Table 1. Effects of microbiome origin on biomass

In general, plant biomass did not differ significantly based on whether the soil microbiome originated from inside or outside of an exclosure. However, responses to changes in microbiome origin were species-specific: the effect was statistically significant for *G. maculatum* ($F_{1, 37} = 6.5450$, $p = 0.01475$) and trended towards significance for *S. flexicaulis* ($F_{1, 37} = 3.4441$, $p = 0.07146$). Biomass when grown in soil microbiomes from within exclosures was approximately 8% higher than outside for *G. maculatum* and 5% higher for *S. flexicaulis*. Microbiome origin also had a statistically significant effect on total native biomass ($F_1 = 5.42$, $p = 0.03177$) but not on total invasive biomass ($F_1 = 0.2017$, $p = 0.6587$).

Overall, these results are generally consistent with prior research. The differing microbiome origins did not have a significant effect on invasive species, while native species tended to be more positively affected when the soil microbiome originated from within an exclosure. This suggests that the detrimental PSFs caused by invasive species outside of exclosures outweigh the harm caused by self-negative PSFs of native species within the exclosures. In addition, it suggests that invasive species are not significantly affected by either the positive or negative PSFs that typically hinder native plants. The effects of microbiome origin also varied with plant life form, consistent with the previously mentioned meta-analysis by Meisner et al. (2014). Microbiome origin had a significant effect on native forbs (*G. maculatum* and *S. flexicaulis*) than the grasses (*C. canadensis* and *M. sinensis*), suggesting that native forbs may be more negatively impacted by invasive PSFs than grasses.

IV. CONCLUSION

This study addresses the previous literature gap on the effects of exclosures and PSFs on the soil microbiome. Future research could focus on analyzing other nature reserves and using a wider variety of species to better understand how the soil microbiomes of exclosures affect native and invasive plant species. In short, these findings can be used to guide conservationists in designing methods to prevent the destruction of native biodiversity and serve as a starting point for further research into the field.

V. ACKNOWLEDGEMENTS

VI. REFERENCES

[1] Van der Putten, W. H., Bradford, M. A., Pernilla Brinkman, E., Voorde, T. F. J., & Veen, G. F. (2016). Where, when and how plant–soil feedback matters in a changing world. Functional Ecology, 30(7), 1109–1121. https://doi.org/10/f8wg5t

[2] Rodriguez, R., & Durán, P. (2020). Natural Holobiome Engineering by Using Native Extreme Microbiome to Counteract the Climate Change Effects. Frontiers in Bioengineering and Biotechnology, 8, 568. https://doi.org/10.3389/fbioe.2020.00568

[3] van der Putten, W. H., Bardgett, R. D., Bever, J. D., Bezemer, T. M., Casper, B. B., Fukami, T., Kardol, P., Klironomos, J. N., Kulmatiski, A., Schweitzer, J. A., Suding, K. N., Van de Voorde, T. F. J., & Wardle, D. A. (2013). Plant-soil feedbacks: The past, the present and future challenges. Journal of Ecology, 101(2), 265–276. https://doi.org/10/f24cwq

[4] Meisner, A., Gera Hol, W. H., de Boer, W., Krumins, J. A., Wardle, D. A., & van der Putten, W. H. (2014). Plant–soil feedbacks of exotic plant species across life forms: A meta-analysis. Biological Invasions, 16(12), 2551–2561. https://doi.org/10/f2484n

[5] Adams, S. N., Jennings, S., & Warnock, N. (2020). Plant invasion depresses native species richness, but control of invasive species does little to restore it. Plant Ecology & Diversity, 13(3–4), 257–266. https://doi.org/10/gj2m44

[6] Gallardo, B., Aldridge, D. C., González-Moreno, P., Pergl, J., Pizarro, M., Pyšek, P., Thuiller, W., Yesson, C., & Vilà, M. (2017). Protected areas offer

refuge from invasive species spreading under climate change. Global Change Biology, 23(12), 5331–5343. https://doi.org/10/cbh8

[7] Van Clef, M. (2014). South Mountain Reservation Forest Regeneration Site Evaluation Report. https://www.somocon.org/wordpress/wp-content/uploads/2015/03/SMR_Forest_Regeneration_Site_Evaluation_Report_2014_12_02.pdf

[8] Howard, M. M., Bell, T. H., & Kao-Kniffin, J. (2017). Soil microbiome transfer method affects microbiome composition, including dominant microorganisms, in a novel environment. FEMS Microbiology Letters, 364(11). https://doi.org/10.1093/femsle/fnx092

X-Net: A Convolutional Neural Network for X-Ray Threat Detection

Ryan Park

Abstract— **This paper proposes X-Net, a novel deep learning architecture that enhances airport security through the detection of dangerous objects in X-ray luggage scans. Scanning of luggage is a critical part of aviation safety but is alarmingly unreliable due to human error. This endangers the safety of millions of airline passengers. To eliminate this human error, several deep learning concepts engineered for the analysis of X-ray baggage scans are introduced. The different concepts are all part of one model, called X-Net. X-Net employs a network of deep convolutional lateral stacks, which combine vertical residual transpose blocks with inter-layer connections. This combination allows for multi-directional gradient flow, resulting in richer and more robust internal feature representations. These innovations enable X-Net to perform exceptionally well on real-world baggage scans, significantly enhancing public safety and potentially saving thousands of lives: X-Net detects malicious items 400% more accurately and 4200% faster than a human TSA officer. Moreover, this proposed approach provides novel and empirically useful deep learning tools that strengthen other fields of computer vision.**

I. BACKGROUND

Now more than ever, airport safety is a high-priority national security issue. Two decades ago, 9/11 exposed the catastrophic results of lackluster aviation security; today, modern terrorism has made the task of protecting airports all the more significant. However, this task is not an easy one, as the ever-increasing rate of air travel has put a strain on security organizations like the TSA. Within the TSA's airport-related responsibilities, X-ray threat detection ("TD") is one of the most important, as it is the primary means of analyzing passengers' luggage. In this high-risk scenario, the TSA has been alarmingly ineffectual. Reports by *Forbes*, *ABC News*, *Newsweek*, and *The New York Times* indicate that the TSA has failed to catch 70-95% of dangerous items in multiple undercover government tests [2], [3], [5], [8]. This astronomical failure rate is no surprise. Ultimately, the failure of baggage screening tests is a result of human error. The logical solution is to remove the human component, i.e., employ automatic threat detection ("ATD") to decrease the failure rate. This paper explores a deep learning model, named X-Net, as an accurate and efficient ATD system to address the issue of X-ray TD.

II. X-NET DESIGN

X-Net combines the YOLO detection head from Redmon et al. (2018) with a novel backbone ("X-Net-backbone"). The detection head is a three-tier residual convolutional network, which detects objects at different input image resolutions [7]. The backbone is the feature extractor, which condenses the information in the input image such that the detection head can then detect objects. This research

introduces "lateral stacks" as the core of X-Net-backbone (inspired by Lin et. al (2017)'s research on feature-pyramid networks). The X-Net-backbone has a main horizontal branch along with several lateral stacks (mini neural
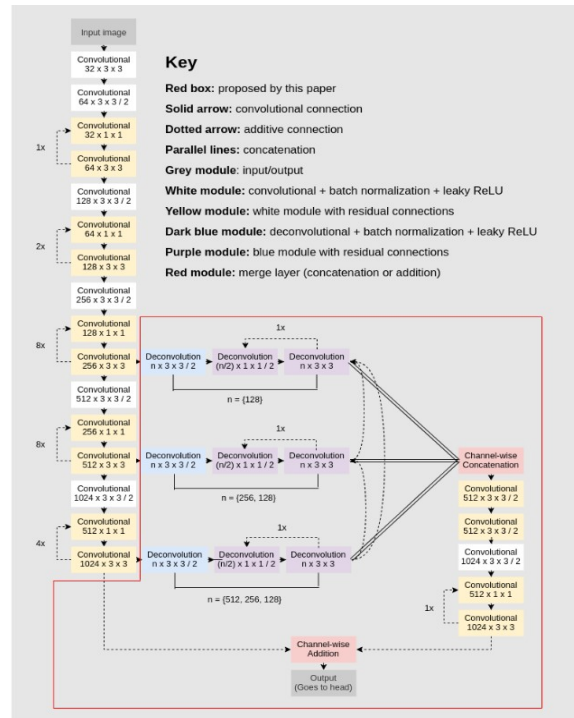
Figure 1. X-Net-backbone. *Deep lateral stacks are purple/blue and are condensed for viewing purposes. Because of space constraints, "vertical" or "lateral" layers appear horizontal, and vice versa.*

networks that stem from the main convolutional branch). The outputs of these stacks are concatenated channel-wise and serve as the input to another network, called the secondary horizontal branch. The output of this branch is then added with the output of the main horizontal branch to form the final output of the X-Net-backbone. See Fig. 1 for a visualization of this architecture, which streamlines information flow and prevents gradient vanishing deep in the network.

III. METHODOLOGY

X-Net is trained and evaluated on SIXray [6]: an X-ray baggage scan dataset created by the Pattern Recognition and Intelligent System Development Laboratory of the University of Chinese Academy of Sciences. SIXray contains 1,059,231 X-ray baggage scans with 12,245 malicious objects, including knives, scissors, guns, wrenches, and pliers. In total, X-Net is trained on ~9,000 X-ray scans with MS-COCO-pretrained weights initialization in early layers. Testing and validation

sets were held-out using ~500 scans each, with no overlap between train/validation/test splits.

X-Net is programmed using Keras and TensorFlow. Open-source code [1] is used for evaluation and results. Code from this study is publicly available at https://github.com/orangese/x-net.

Training utilizes a GeForce RTX 2060 for 72 hours. Due to hardware limitations, a resource-conscience training schema is employed. X-Net is trained in three different stages, using Adam with a batch size of 1 and learning rates = [0.001, 0.0001, 0.00001]. This annealed training strategy initially encourages exploration of parameter space (learning rate = 0.001) and then encourages convergence towards a minimum of the loss function (learning rate = 0.00001).

IV. RESULTS AND DISCUSSION

The main innovation of X-Net are the aforementioned "deep lateral stacks": these are mini-networks that branch off of several points within the convolutional backbone of X-Net. Unlike in previous research, which used largely linear network structures like the Inception and ResNet families [6], X-Net incorporates information across several resolution "checkpoints" along the backbone. Thus, the detection of smaller malicious objects is emphasized compared to previous convolutional networks that lossily compress tensor dimensions deep into the network.

Two evaluation metrics from [6] are used to evaluate X-Net: classification accuracy (measures detection of malicious items) and localization accuracy (measures the ability to identify the location of malicious items). In this section, accuracy is computed using mean average precision ("mAP"), a standard computer vision metric used to assess object detection algorithms. All reported percent results are relative percent increases, not absolute increases in mAP.

With respect to classification accuracy, X-Net outperforms the next-best ATD model by 9.93% and plain YOLOv3 by 11.03% (Table 1):

TABLE I.   CLASSIFICATION MAP (%) OVER ALL CLASSES AND MODELS

Method	Item Category					
	Gun	Knife	Wrench	Pliers	Scissors	Overall
ResNet-101 [6]	87.65	84.26	69.33	85.29	60.39	77.38
ResNet-101+CHR [6]	85.45	87.21	71.23	88.28	64.68	79.37
Inception-v3 [6]	90.05	83.80	68.11	84.45	58.66	77.01
Inception-v3+CHR [6]	88.90	87.23	69.47	86.37	65.50	79.49
YOLOv3	94.65	77.75	71.03	83.10	66.96	78.70
X-Net	96.28	83.52	78.71	88.16	90.24	87.38

More importantly, X-Net decreases human error by several orders of magnitude, achieving a 399.31% gain in classification mAP and a 4247.83% gain in speed over a TSA officer (Fig. 2):
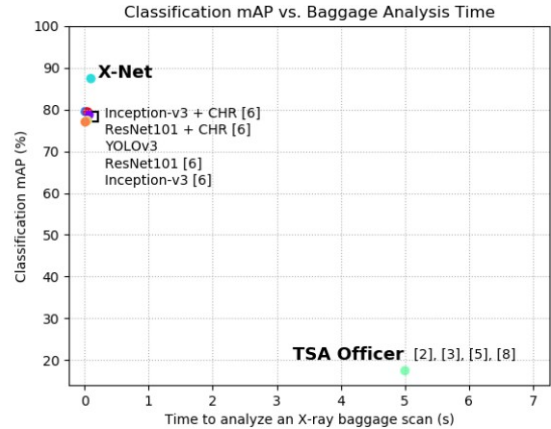

Figure 2. Classification mAP (%) vs. Baggage Analysis Time (s).

With respect to localization accuracy, X-Net improves overall mAP by 17.93% (Table 2):

TABLE II.   LOCALIZATION MAP (%) OVER ALL CLASSES AND MODELS

Method	Item Category					
	Gun	Knife	Wrench	Pliers	Scissors	Overall
ResNet-101 [6]	73.77	65.13	28.34	62.24	21.02	50.10
ResNet-101+CHR [6]	80.86	73.85	52.41	9.30	40.34	51.35
Inception-v3 [6]	79.94	75.38	59.36	59.58	40.34	62.92
Inception-v3+CHR [6]	78.70	74.36	52.41	59.96	52.27	63.54
YOLOv3	78.35	55.11	43.68	51.90	39.39	53.68
X-Net	87.95	72.01	67.24	73.90	73.55	74.93

V. CONCLUSION

Ultimately, X-Net advances both deep learning theory and its applications. First, it provides empirical justification for the proposed lateral stacks. Specifically, X-Net increases detection and localization of the smallest class (scissors) by 34.77% and 40.71%, implying that lateral stacks are especially useful for detecting hard-to-see objects. Similar accuracy gains can be noted for the other objects in Tables 1 and 2. However, the accuracy gains for larger objects (guns, knives, wrenches, etc) were smaller. This result is consistent with the intended design of X-Net: unlike previous networks, X-Net is able to access information at various stages within the network (i.e., both high-resolution and low-resolution) due to its deep lateral stack paradigm. As a result, it can detect smaller objects with relative ease. Increasing detection accuracy for the larger objects, however, is a potential extension of this research that will likely be addressed as convolutional network design continues to evolve.

X-Net outperforms the TSA by 399.31% in terms of malicious item detection accuracy and is incredibly fast, achieving a 4247.83% reduction in time per scan. Therefore, X-Net is a critical enhancement of airport security and a practical solution to the problem of X-ray TD. If integrated into security systems, it would detect many threats and save many lives, thereby helping to preserve public safety

VI.   REFERENCES

[1] Cartucho. (2018, March 18). Retrieved from https://github.com/Cartucho/mA

[2] Goldstein, M. (2017, November 9). TSA Misses 70% Of Fake Weapons But That's An Improvement. *Forbes*. Retrieved from

   https://www.forbes.com/sites/michaelgoldstein/2017/11/09/tsa-misses- 70-of-fake-weapons-but-thats-an-improvement/#1d42a8b32a38

[3] Kerley, D., & Cook, J. (2017, November 9). TSA fails most tests in latest undercover operation at US airports. *ABC News*. Retrieved from https://abcnews.go.com/US/tsa-fails-tests-latest-undercover-operation-us-

   airports/story?id=51022188

[4] Lin, T.-Y., Dollar, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. (2017). Feature Pyramid Networks for Object Detection. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). doi: 10.1109/cvpr.2017.106

[5] Meza, S. (2017, October 11). TSA Fails to Spot Weapons More than Half the Time. *Newsweek*. Retrieved from https://www.newsweek.com/tsa-failshalf-time-706568

[6] Miao, C., Xi, L., Wan, F., Su, C., Liu, H., Jiao, J., & Ye, Q. (2019). SIXray: A Large-scale Security Inspection X-ray Benchmark for Prohibited Item Discovery in Overlapping Images. Retrieved from https://arxiv.org/pdf/1901.00303.pdf

[7] Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement. Retrieved from https://pjreddie.com/media/files/papers/YOLOv3.pdf

[8] Smith, J. F. (2015, June 2). Head of T.S.A. Out After Tests Reveal Flaws. *The New York Times*. Retrieved from https://www.nytimes.com/2015/06/02/us/head-of-tsa-out-after-tests-reveal-

   flaws.html

Genetic Association Testing and Predictive Modeling for Non-Small Cell Lung Carcinoma RNA Sequencing

Lyon Kim

Abstract— **We developed Python code to analyze a bulk RNA sequencing dataset consisting of lung tissues from healthy and non-small cell lung carcinoma (NSCLC) patients. Our preliminary goal was to find genes that were positively or negatively associated with cancer. To do so, we tested the null hypothesis, which stated that the average gene expression between cancer and non-cancer patients is equal. We rejected the null hypothesis since 86.8% of the genes showed a difference in expression. After finding genes associated with cancer, we built a machine learning logistic regression model from the gene expression data. To efficiently measure the performance of our method's ability to predict cancer, we randomly split the data into training (80% of data) and testing datasets (20%) and used five-fold cross validation. By adjusting the probability threshold for classifying cancer, we created an ROC curve, representing the trade-off between the fpr (false positive rate) and the tpr (true positive rate). Ultimately, we hope that these mechanisms can help increase the chance for an early diagnosis of NSCLC, which is crucial to controlling and even preventing it.**

I. INTRODUCTION

X = gene expression (matrix - $N \times M$)
N = number of patients (1118)
M = number of genes (10077)
y = cancer labels (vector of length N) (0 is non-cancer, 1 is cancer)
μ^n = vector of means of each gene in non-cancer patients (has length of M)
μ^c = vector of means of each gene in cancer patients (has length of M)
N_n = number of non-cancer patients
N_c = number of cancer patients
σ_n = vector of the standard deviation of the gene expression of non-cancer patients
σ_c = vector of the standard deviation of the gene expression of cancer patients
X_{ji}^n = an arbitrary data point from a non-cancer patient that is drawn from a normal distribution in which the mean is μ^n and the variance is σ_n^2
X_{ji}^c = an arbitrary data point from a cancer patient is drawn from a normal distribution in which the mean is μ^c and the variance is σ_c^2
Z = Z values: represent how many standard deviations are in the difference of empirical means (has length of M)
$\bar{Z}$ = normalized Z values: Z values scaled to standard normal distribution under the null hypothesis
ℓ_j = the quantitative measure of how much each patient corresponds to the likelihood of cancer (has length of N)
p_j = predicted probability that patient j has cancer
β = vector of coefficients for every gene that directly relates to getting the probability of cancer
D^c = a measure of the amount of cancer genes expressed by the patient
D^n = a measure of the amount of non-cancer genes expressed by the patient

We assumed that each data point (X_{ji}^n and X_{ji}^c) is normally distributed:

$$X_{ji}^n \sim \mathcal{N}(\mu_i^n, (\sigma_i^n)^2) \quad X_{ji}^c \sim \mathcal{N}(\mu_i^c, (\sigma_i^c)^2) \quad (1)$$

First, we calculate the gene expression means ($\hat{\mu}^n$ and $\hat{\mu}^c$) for each gene i for $1 \leq i \leq M$

$$\hat{\mu}_i^n = \frac{\sum_{j=1}^{N_n} X_{ji}^n}{N_n}, \quad \hat{\mu}_i^c = \frac{\sum_{j=1}^{N_c} X_{ji}^c}{N_c} \quad (2)$$

We calculate σ_n and σ_c

$$(\sigma_i^n)^2 = \frac{\sum_{j=1}^{N_n}(X_{ji}^n - \mu_i^n)^2}{N_n}, \quad (\sigma_i^c)^2 = \frac{\sum_{j=1}^{N_c}(X_{ji}^c - \mu_i^c)^2}{N_c} \quad (3)$$

Figure 1. Defining Notation

II. METHOD

Discovering Genes Associated with NSCLC

We aimed to detect genes with a significant difference in gene expression between cancer and non-cancer patients. To do so, we tested the null hypothesis that the average gene expression in both types of patients is equal. The data we used was from a normalized, merged lung cancer transcriptome dataset that listed for each patient whether each gene was expressed or not through methods of differential expression analysis, batch effect correction, and filtering of genes with low variance.

Testing each Gene's Significance

Next, we attempted to test the null hypothesis that the two means are equal ($\mu^n = \mu^c$). For genes that we reject the null hypothesis, it is unlikely that differences between the two conditions are solely due to experimental noise. The empirical average of the sample for non-cancer and cancer patients can be calculated as:

$$\hat{\mu_i}^n \sim \mathcal{N}(\mu_i^n, \frac{(\sigma_i^n)^2}{N_n}), \quad \hat{\mu_i}^c \sim \mathcal{N}(\mu_i^c, \frac{(\sigma_i^c)^2}{N_c}) \quad (4)$$

Next, we define σ^2 to be a measure of the variability of each gene, Z, and $\bar{Z}$. The Z value represents the number of standard deviations between the difference of cancer and non-cancer means for a given gene.

$$\sigma_i^2 = \frac{\frac{(\sigma_i^n)^2}{N_n} + \frac{(\sigma_i^c)^2}{N_c}}{1/N_n + 1/N_c} \quad Z_i = \frac{\hat{\mu_i}^c - \hat{\mu_i}^c}{\sigma_i} \quad \bar{Z}_i = \frac{\hat{\mu_i}^c - \hat{\mu_i}^c}{\sigma_i\sqrt{\frac{1}{N_n} + \frac{1}{N_c}}} \quad (5)$$

We can then calculate the null distribution of Z.

$$Z_i \sim \mathcal{N}(0, \frac{1}{N_c} + \frac{1}{N_n}), \quad \bar{Z}_i \sim \mathcal{N}(0, 1) \quad (6)$$

Now, in order to test the null hypothesis, we tested whether $|\bar{Z}| > 3.33$ and found 8747 genes to be significant. Next, we prove that genes passing the null hypothesis test have a p-value of less than 0.001.

$$P(|\bar{Z}|_i > 3.33) = 2(P(\bar{Z}_i > 3.33)) \approx 0.00086 < 0.001 \quad (7)$$

Figure 2. Method for Testing Gene Significance

Creating a Logistic Regression Model for the Prediction of NSCLC

In the second part of our project, we used the package scikit-learn to fit a logistic regression model to our data. By five-fold cross validating our model, we were able to show that our model predicts cancer with high accuracy from gene expression.

In the logistic regression model, we predict the probability of having cancer (p_j) as:

$$p_j = \frac{1}{1 + e^{-\ell_j}} \quad (8)$$

Now we define ℓ as written below with β_0 being the intercept:

$$\ell_j = \sum_{i=1}^{M}(\beta_i * X_{ji}) + \beta_0 \quad (9)$$

We define the likelihood (L) as the probability of observing y given X, which is the product of the probabilities of observing each patient's condition. β is chosen to maximize the likelihood of observing all the data points.

$$L = P(y|X) = \prod_{j=1}^{N}(p_j^{y_j} * (1 - p_j)^{1-y_j}) \quad (10)$$

Next, we randomly split our data into training (80%) and testing data (20%), with a different 20% for each of the five partitions. This method is used to measure our ability to predict future data based on past data. In order to classify patient j as cancer or non-cancer, we thresholded the predicted probabilities at α. By varying α from 0.00 to 1.00, we generated an ROC curve of the true positive rate (tpr) vs. the false positive rate (fpr). The tpr is defined as proportion of cancer patients correctly classified. The fpr, on the other hand, is the proportion of non-cancer patients incorrectly classified.
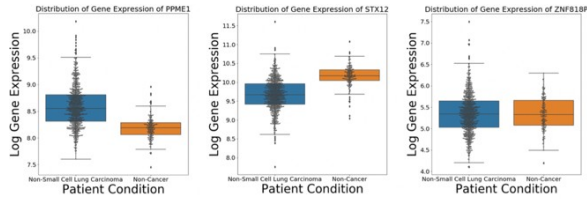
18

III. RESULTS

Figure 3. Genes Associated with NSCLC

These three graphs are boxplots of gene expressions in both cancer and non-cancer patients.

- Left: This boxplot represents the gene, PPME1, with the largest decrease in average gene expression from cancer to non-cancer patients.
- Middle: This second graph represents the gene, STX12, with the largest increase in average gene expression from cancer to non-cancer patients.
- Right: This gene, ZNF818P, was found to have no change in average gene expression from cancer to non-cancer patients. This graph serves as a control gene.

Therefore, the gene PPME1, which is highly prevalent in cancer patients, is associated with NSCLC, whereas the gene STX12 is negatively associated with NSCLC. Studying these genes may yield insight into the biological mechanism of NSCLC. In fact, inhibition of PPME1 has already been proven to prevent cell proliferation and induce apoptosis in certain cancer cells [2]. On the other hand, the gene STX12, which is a part of the SNARE complex, helps to control tumorigenesis by regulating cancer cell invasion [3].
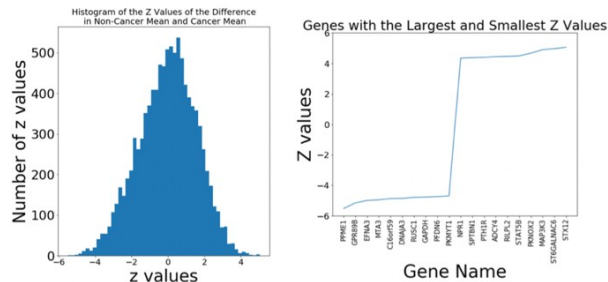

Figure 4. Genes with the Largest and Smallest Z Values are most associated with NSCLC

By finding the Z values, we were able to quantify the difference of the average gene expression between cancer and non-cancer patients. We used the Z values to find candidate genes that could be related to cancer. Next, we tested for each gene for significance, and we found 8747 significant genes, including the twenty genes in the figure on the left.
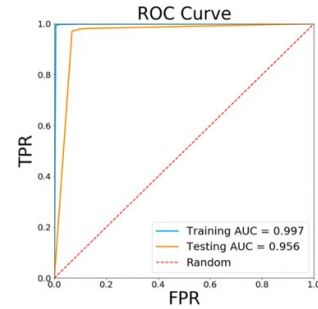

Figure 6. Training and Testing Accuracy

We randomly split our data into training (80%) and testing (20%) data in five fold cross-validation. This method is used to measure our ability to predict future data based on past data. We use the training data to create the model and use the testing set as if it were new data to see how well our model works. The ROC curve model performed significantly better on the training data (Fig. 6). On the other hand, the testing data represents a more realistic measure of an ability to predict on new patients. The overall testing accuracy was around 96%, indicating that this model can be used for early detection of NSCLC. With methods for early diagnosis, it could lead to an improved prognosis. This model can be further improved in the future by training the model with more data and more careful gene selection.

IV. CONCLUSION

In this project, we successfully detected more than 20 genes that were both positively and negatively associated with cancer, specifically non-small cell lung carcinoma (NSCLC), and fit a logistic regression model that predicts a patient's risk of cancer based on their RNA sequencing data. In the future, with the provided merged lung cancer transcriptome dataset, we would modify our predictive logistic regression model to take in only the genes that we found to be significant, or associated with cancer, hence combining my two methods. We would also test for significant genes at different stages of cancer. From this, we could predict a patient's risk of a certain stage of NSCLC and their survival rate. We believe that by finding genes associated with cancer, we can shed light on the biological mechanism of cancer and raise awareness about these genes. Our predictive logistic regression model can be used for early detection of NSCLC. In addition, we believe that it can also be used to detect other types of cancer by training a new logistic regression model on another dataset.

V. REFERENCES

[1] Lim, Su Bin, Swee Jin Tan, Wan-Teck Lim, and Chwee Teck Lim. "A merged lung cancer transcriptome dataset for clinical predictive modeling." Scientific data 5 (2018): 180136.
[2] Li, Jing, Sufang Han, Ziliang Qian, Xinying Su, Shuqiong Fan, JiangangFu, Yuanjie Liu et al. "Genetic amplification of PPME1 in gastric and lung cancer and its potential as a novel therapeutic target." Cancer biology & therapy 15, no. 1 (2014): 128-134.
[3] Meng, Jianghui, and Jiafu Wang. "Role of SNARE proteins in tumourigenesis and their potential as targets for novel anti-cancer therapeutics." Biochimica et Biophysica Acta (BBA)-Reviews on Cancer 1856, no. 1 (2015): 1-12.

DeepNeuroNet: A Novel Multiclass Model to classify Brain Tumors and Neurodegenerative Diseases using Machine Learning

Vaibhav Mishra

Abstract— **One of the most important applications of machine learning is the use of deep learning models in medical diagnosis and treatment. This application is particularly valuable in the early diagnosis of brain tumors and neurodegenerative diseases because earlier intervention leads to better prognosis and prevention of more fatal conditions. Among the most widely affected brain disorders are brain tumors, Alzheimer's disease (AD), and Mild Cognitive Impairment (MCI). These diseases have a high incidence of 11.3% and a high mortality rate of 40% [6]. Therefore, this study aimed to help with the early diagnosis of these diseases through the development of a new multiclass convolutional neural network (CNN) model to classify glioma tumors, meningioma tumors, pituitary tumors, AD, and MCI from normal patients with an overall accuracy of 88.33%. DeepNeuroNet was the first model that used brain MRIs to classify both brain tumors and neurodegenerative diseases. This model would have many applications including brain tumor detection and possible treatment research in clinical settings and the potential to be used for the early diagnosis of brain diseases.**

I. INTRODUCTION

Each year, millions of people worldwide are diagnosed with brain tumors, Alzheimer's disease (AD), and Mild Cognitive Impairment (MCI) [6]. These diseases can cause gradually worsening motor and cognitive functioning capabilities in patients and can ultimately result in death. Doctors and radiologists seek early diagnosis of these diseases in order to treat patients with the medications while the disease is still in mild stages to slow down progression. Before deep learning models, classical imaging techniques had limited accuracy and capabilities to classify multiple diseases. Furthermore, the previous techniques had longer runtimes and were often misdiagnosed at a higher rate. With the advances in machine learning, healthcare practitioners can now use deep learning models to fasten the diagnosis of these brain diseases. The advantage of deep learning models over traditional methods is that deep learning models are more accurate, faster, cost-efficient, and easier to implement in diagnosing diseases. This enables deep learning models to be able to detect even subtle brain abnormalities in MRI images faster and more accurately increasing the likelihood of early diagnosis.

Recently, there have been new models developed to help with the early diagnosis of brain tumors. For instance, Irmak proposed a new multiclass model that classified brain tumors by location in the brain and the specific neural cell affected. The study focused on diagnosing glioma, meningioma, pituitary, and metastatic tumors from healthy patients with a very high accuracy of 92.66% [2]. The study also showed that the new convolutional neural network (CNN) model had a higher accuracy in classifying the tumors than VGG16 or ResNet. Another study by Marghalani and Arif developed a 3-class model that classified AD, brain tumors, and healthy normal patients (unaffected by any neurological disorder)

with an average accuracy of 97% [4]. There have also been developments in differentiating between AD and MCI from normal patients. Basaia et.al developed a deep learning algorithm that achieved a 75.4% accuracy for AD versus cMCI classification [5].

Even though there have been multiclass models that classify brain tumors as well as neurodegenerative diseases, there was no model that combined the two models into a 6-class model with a high accuracy, which would be highly useful in clinical diagnosis because of a wider range of disease classes.

Therefore, the purpose of this study was to create a 6-class CNN model which could detect glioma tumor, meningioma tumor, pituitary tumor, AD, and MCI. This would be the first such model to combine the classification of the different types of brain tumors as well as two of the most common neurodegenerative diseases. This model would have great potential in the early diagnosis of patients.

II. METHODS

The images for the brain tumors and normal patient MRIs were collected from the Kaggle database [1] and the images for the AD and MCI images were taken from the Alzheimer's Disease Neuroimaging Initiative (ADNI) [3]. In total, 200 images were collected for each disease and split into training (60%), validation (30%), and test subsets (10%). The number of images by disease class is shown below in table 1.

TABLE 1: NUMBER OF MRI IMAGES BY DISEASE CLASS

Data Folder	Disease Class					
	AD	MCI	Glioma	Pituitary	Normal	Meningioma
Training	120	120	120	120	120	120
Validation	60	60	60	60	60	60
Test	20	20	20	20	20	20

The images were scaled and normalized to improve accuracy. Data augmentation was applied to all the images including sharing, zooming, and flipping them. The images were trained in the DeepNeuroNet model for 20 epochs with the Adam optimizer and the categorical cross entropy loss. The DeepNeuroNet model consisted of 12 layers including convolutional layers, max-pooling layers, dropout layers, and batch normalization layers followed by a flatten and a dense layer. The total number of parameters was 346,374.
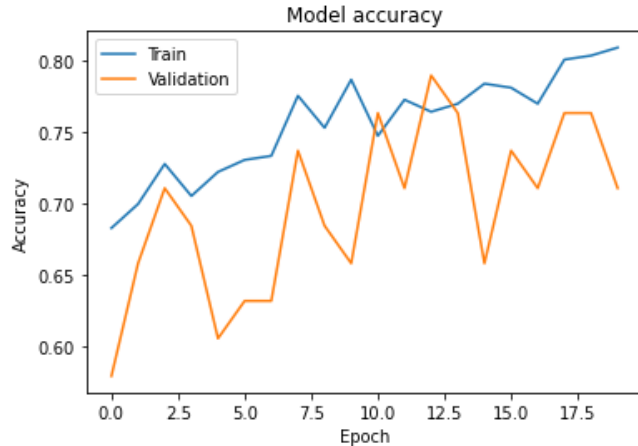
III. RESULTS

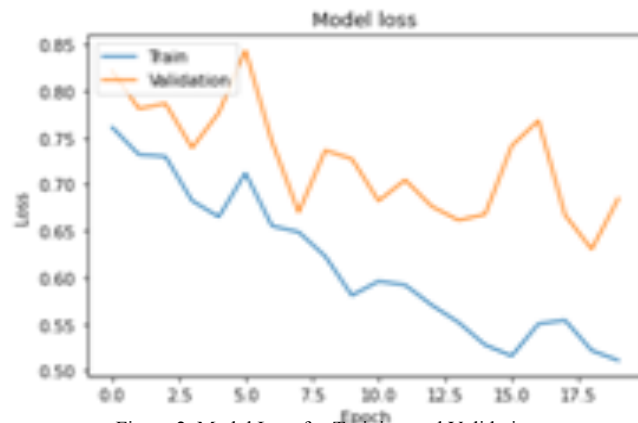Figure 1: Model Accuracy for Training and Validation

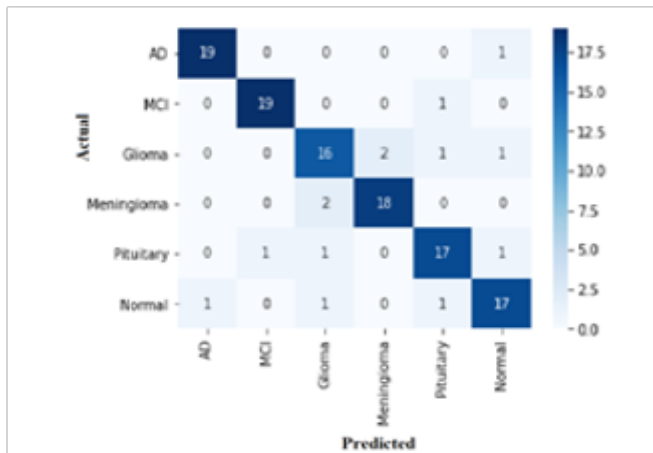Figure 2. Model Loss for Training and Validation

Figure 3. Confusion Matrix for DeepNeuroNet

After running the DeepNeuroNet model for 20 epochs, the model accuracy and loss graphs were generated as shown in Figure 1 and Figure 2 respectively . Then, the training set was run on the model. The model correctly classified 106 images from the original 120 MRI images as shown in Figure 3. Thus, the overall test accuracy was 88.33%, a relatively high accuracy for a 6-class model [4].
The classification report on the precision, recall, support, and

f1-scores of each of the 6 classes was also generated (Table I) . The model recognized AD and MCI the best with a 95% precision for both classes. This could be a result of the more prevalent features of AD and MCI compared to the 3 classes of tumors which were hard to distinguish from each other.

	Precision	Recall	F1-Score	Support
AD	95.00%	95.00%	95.00%	20
MCI	95.00%	95.00%	95.00%	20
Gioma	80.00%	80.00%	80.00%	20
Meningioma	90.00%	90.00%	90.00%	20
Pituitary	85.00%	85.00%	85.00%	20
Normal	85.00%	85.00%	85.00%	20

Figure 3: Confusion Matrix for DeepNeuroNet

Overall, the model performed well and achieved high training and testing accuracy with fewer parameters than many other studies meaning that the model was quicker and more efficient in classifying more diseases. Although there have been many multiclass models for brain diseases like a 4-class model developed by Singh et.al [6], the DeepNeuroNet model proposed in this study was the first model to conduct a 6-class diagnosis of neurological disorders with such a high accuracy.

IV. CONCLUSION

The DeepNeuroNet model was proposed and shown to achieve a high accuracy in classifying glioma tumors, meningioma tumors, pituitary tumors, AD, and MCI from normal patients making it the first model to use a novel CNN architecture to achieve a 6-class diagnosis. This model has the potential to help with the early diagnosis of brain tumors and neurodegenerative diseases and therefore, save millions of lives worldwide. Further studies could enhance this model by adding more classes of diseases including PSP, MSA, and CBD as well as using non imaging data like gene expression to help the model be able to classify at a higher accuracy through a wider range of data.

V. REFERENCES

1.   Bhuvaji, S., Kadam, A., Bhumkar, P., Dedge, S., & Kanchan, S. (2020). *Brain Tumor Classification (MRI]) [Data set]*. Kaggle. https://doi.org/10.34740/KAGGLE/DSV/1183165.
2.   Irmak, E. (2021, April 22). *Multi-classification of brain tumor MRI images using deep convolutional neural network with fully optimized framework*. Iranian Journal of Science and Technology, Transactions of Electrical Engineering. Retrieved November 20, 2021, from https://link.springer.com/article/10.1007/s40998-021-00426-9.
3.   Petersen, R. C., Aisen, P. S., Beckett, L. A., Donohue, M. C., Gamst, A. C., Harvey, D. J., Jack, C. R., Jagust, W. J., Shaw, L. M., Toga, A. W., Trojanowski , J. Q., & Weiner , M. W. (2009, December 30). *Alzheimer's*

disease neuroimaging initiative (ADNI): Clinical characterization. Neurology. Retrieved November 20, 2021, from https://pubmed.ncbi.nlm.nih.gov/20042704/.

4.　　Marghalani, B. F., & Arif, M. (2019). *Automatic Classification of Brain Tumor and Alzheimer's Disease in MRI.* Procedia Computer Science, vol. 163, pp. 78-84, 2019.

5.　　Basaia, S., Agosta, F., Wagner, L., Canu, E., Magnani, G., Santangelo, R., & Filippi, M. (2018, December 18). *Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and Deep Neural Networks.* NeuroImage. Clinical. Retrieved November 20, 2021, from https://pubmed.ncbi.nlm.nih.gov/30584016/.

6.　　Singh, G., Vadera, M., Samavedham, L., & Lim, E. C.-H. (2019, May 12). *Multiclass diagnosis of neurodegenerative diseases: A neuroimaging machine-learning-based approach†.* ACS Publications. Retrieved November 20, 2021, from https://pubs.acs.org/doi/abs/10.1021/acs.iecr.8b06064.

Enhancing Efficacy and Cost Effectiveness of Air Filtration Systems by Optimized Nanoparticle Deposition

Ishika Nag

Abstract— **Every year, seven million people die from severe cardiovascular and respiratory diseases, caused by ambient and household air pollution. An increase in air pollution from particulate matter less than 2.5 microns in diameter (PM$_{2.5}$) has shown to be a significant contributor of cardiovascular and respiratory diseases. The goal of this study was to develop an efficient and cost-effective air-filtration system by the impregnation of selected nanoparticles, utilizing their high surface-to-volume ratio to entrap PM$_{2.5}$. The experimental set-up consisted of a wind tunnel with incense sticks as the particulate matter source, measured by laser particle detectors upstream and downstream of the filters. Results found that a mixture of zinc oxide, titanium dioxide & graphene improved filtration efficiency of a baseline filter by 206%. There was also a 70% improvement in the cost of the filters. The versatility and cost-effectiveness of this design makes it applicable for personal masks & filters, air-conditioning filters, car-cabin filters, and fire-fighting equipment. The correlation between air pollution and fatalities from viral infections suggests that abatement technologies with innovative filtration systems are critical in saving human lives.**

I. INTRODUCTION

Seven million people die every year because of air pollution [1]. The majority (91%) of the world's population lives in locations exceeding the World Health Organization's air quality guidelines. Due to its small size, PM$_{2.5}$, particulate matter less than 2.5 microns in diameter, is capable of penetrating deep into lung passageways and entering the bloodstream, causing and aggravating cardio-vascular, cerebro-vascular, and other respiratory diseases [2]. Furthermore, long-term exposure to air pollution has been found to increase the vulnerability of contracting COVID-19 [3]. An increase of only 1 μg/m^3 in PM$_{2.5}$ is associated with an 8% increase in the COVID-19 death rate in the United States [3]. Abatement technologies such as ionic and High Efficiency Particulate Air (HEPA) air filtration systems [4] have been developed to filter PM$_{2.5}$ particles,but remain quite expensive and hence unaffordable to communities with limited resources [5, 6]. Therefore, a cost-effective and efficient abatement system is essential to help resolve the issue.

Nanoparticles have a high surface-to-volume ratio, which enhances the entrapment of particulate matter by adsorption. The surface adsorption energy is unique to the small size of nanoparticles with extremely high surface to volume ratios, where the unsaturated surface chemical bonds tend to adsorb other chemicals or biomolecules to reduce their surface energy [7]. The three nanoparticles used for this study were graphene, titanium dioxide (TiO$_2$), and zinc oxide (ZnO), which have

been known to have filtration properties due to their high adsorption capabilities [8, 9, 10]. Graphene, an allotrope of carbon consisting of a single layer of carbon atoms arranged in a hexagonal lattice structure, has high adsorption capacities mainly due to these unique nanostructures, and hence have been proven to be efficient in the capture of particulate matter [11, 12].

The current work is aimed to develop an efficient and cost-effective air-filtration system by an optimized deposition of nanoparticles, based on their air filtration capabilities, clinical safety, and non-toxicity. The filtration system needs to be versatile and effective at different pollution levels in different parts of the world. The goal of this work is to also develop a simple application technique of the nanoparticles such that it can be easily applied to various filtration systems, thus providing an affordable alternative to expensive high quality air filtration devices with comparable air filtration capabilities.

II. METHOD

The nanoparticles (NPs) used for this study were titanium oxide (TiO), zinc oxide (ZnO) and graphene. The combination of NPs was mixed with ethanol to create a suspension. This was then aerosolized and sprayed on to the air filters, using the pressurized sprayer system. A 'high-quality' air filter (MERV-14, FPR-10, HEPA) was used for comparison and benchmarking the filtration efficiency in order to validate the results from this experiment, compared to previously performed studies [13]. The deposition method of NPs onto the filtration media was also varied and tested for uniform spatial distribution. Different spray mechanisms were tested using pipettes, spray bottles and pressurized sprayers. Different zones of the air filter, in 9 locations, were tested for spatial consistency in filtration efficiency (Fig. 1). The pressurized spray application resulted in the most uniform spatial distribution of the NPs, and it was chosen as the preferred application method for its simplicity and
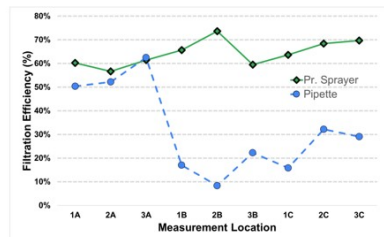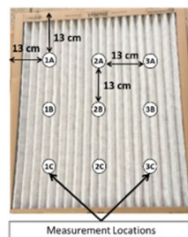
effectiveness.

Figure 1. The pressurized sprayer system demonstrated better nanoparticle deposition uniformity, verified with the spatial consistency test.

The surface morphology of the filters was characterized using the scanning electron microscope (SEM) imaging technique, confirming the adhesion of nanoparticles to the filters in the 'before' images and the entrapment of particulate matter onto the nanoparticle surfaces in the 'after' images of the different nanoparticle coated filters (Fig. 2).
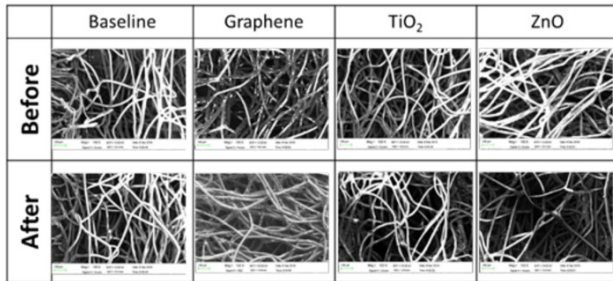

Figure 2. Scanning Electron Microscope (SEM) images of uncoated and coated filters before and after exposure to PM particles, confirm the adhesion of NPs and entrapment of PM.

A full-factorial Design of Experiments (DOE) statistical analysis model was used to randomize the run order of the experiment, minimize bias, and aid with the Analysis of Variance (ANOVA) study. A statistical repeatability and reproducibility study (Gage R&R) was used to determine the measurement uncertainty of the experiment. The tests were repeated for 10 trials each. As seen in Fig. 3, 95% of the contribution was from 'part-to-part variation' and 5% from the process variation.
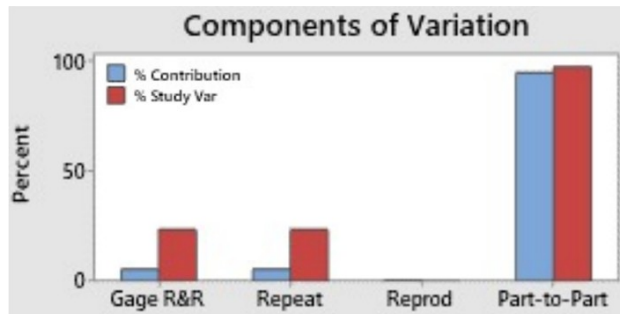

Figure 3. The Gage Repeatability and Reproducibility analysis of the experimental set-up indicating acceptable levels of measurement uncertainty

III. EXPERIMENTAL DESIGN

A wind tunnel was designed and created to test for the efficiency of the filters (Fig. 4). Incense sticks were used as the source of PM and laser particle detectors measured the PM at the inlet and outlet sections of the wind tunnel. A manometer was used to measure pressure drop, and a lamp was used on the TiO-coated filters for photocatalysis testing.
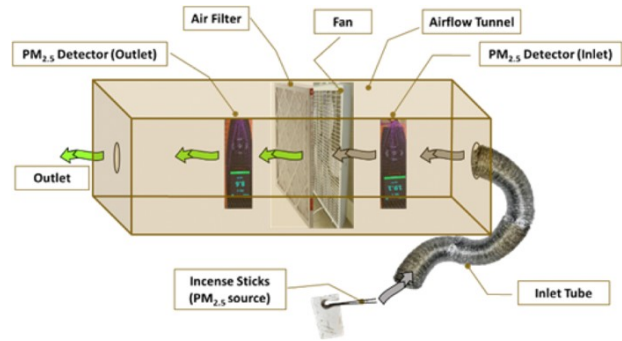
Figure. 4. The experimental set-up included a wind tunnel with a fan, incense sticks to simulate the PM and laser particle detectors to measure filtration efficiency.

IV. RESULTS AND DISCUSSION

The results showed (Fig. 5) that the concentration of NPs has a direct correlation to the filtration efficiency (1). The correlation between nanoparticle type and filtration efficiency was also observed, with TiO_2 coated filters demonstrating the highest efficiency.
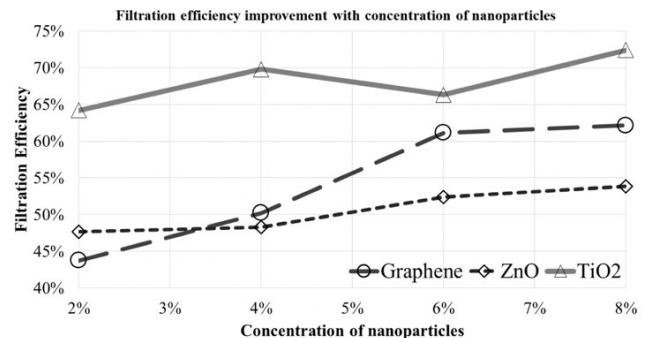
Figure. 5. Increase in nanoparticle concentration on coatings improves filtration efficiency

The TiO_2 coated filters had a 7% increase in filtration efficiency when placed under light. TiO_2, with its photocatalytic properties, absorbs the ultraviolet component of sunlight which excites the electrons from its valence band to the conduction band. They act as a catalyst to form the superoxide anion ($O•^-_2$) and reactive hydroxyl (OH•) radicals from atmospheric moisture and oxygen. They are then able to react with the $PM_{2.5}$ particles converting them into CO_2 and H_2O [14].

A significant parameter that affects the filtration system's energy consumption is its pressure drop. Pressure drop was

calculated by using the Bernoulli equation, which is given by (Eq. 1):

$$Filtration\ Efficiency\ (\%) = \frac{PM_{2.5\,inlet} - PM_{2.5\,outlet}}{PM_{2.5\,inlet}} \times 100$$

. Assuming a steady, incompressible, and frictionless flow along a streamline, with the same horizontal height; this can be simplified to the pressure drop equation (2):

$$\Delta P = \frac{1}{2}\rho(v_1^2 - v_2^2)$$

A quantitative test of airflow, conducted by measuring the pressure drop (Δp) across the filter, determined that the application of NPs to air filters does not affect their energy consumption in a measurable way.

The NP coatings consistently demonstrated the ability to improve the filtration efficiency of a baseline filter (Fig. 6). The filter, coated with a mixture of the three-NPs, had the highest filtration efficiency which was 206% higher than a baseline filter. The filtration efficiency of this filter, at 77%, was quite comparable to the more expensive 'high-quality' FPR10 filters. The filter coated with TiO_2 alone was also quite effective, but less versatile due to its dependence on light for activation of its photocatalytic properties.
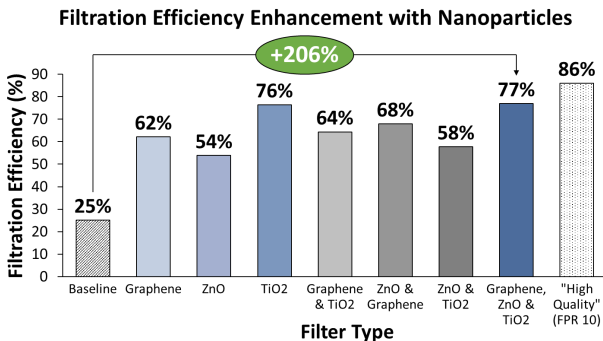

Fig. 6 The Gr/ZnO/TiO2 combination is 206% more effective than a baseline uncoated filter and is almost similar to a 'high-quality' HEPA filter

An accelerated durability testing was also performed to test the effectiveness of the filters over longer usage periods. 90% of the filter's effectiveness was maintained after 50 equivalent days of operation.

Considering the baseline cost of a commercially available filter and the additional cost of NPs and processing, the nanoparticle coated filters were 70% cheaper than the HEPA filters, and 99% cheaper than ionic filters (Fig. 7).
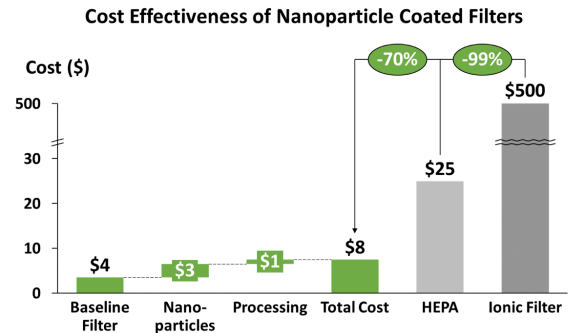
Fig. 7. NP coated filters are significantly less expensive than HEPA or ionic filtration systems, while having comparable filtration efficiency.

V. CONCLUSION

The nanoparticle coatings consistently demonstrated the ability to improve the filtration efficiency of a baseline filter. The mixture of the three nanoparticles - graphene, TiO_2 and ZnO - improved the efficiency of a baseline filter by 206%. The TiO_2 coated filter was also quite effective and demonstrated its photocatalytic effectivity, but less versatile due to its dependence on light. Cost-effectiveness was one of the main objectives of this experiment in order to make this technology available to societies with limited resources. Considering the baseline cost of a commercially available filter and the additional cost of nanoparticles and processing, the nanoparticle coated filters were 70% cheaper than the HEPA filters, and 99% cheaper than ionic filters. The safety of nanoparticle usage is of utmost importance and continues to be a subject of research worldwide[15, 16]. The nanoparticles chosen for this study are known for their clinical safety and non-toxicity and are extensively used in cosmetic and biomedical applications, e.g., pill coatings, sunscreens[17]. These NP coatings can be used in several applications including face masks, air-conditioning and car cabin filters, fire-fighting masks, and industrial pollution control systems. The versatility and effectiveness of this nanoparticle coated filtration system makes it applicable for varying pollution levels in different parts of the world. There is a significant correlation between air pollution and deaths from respiratory diseases and virus infections like COVID-19 [3]. Such novel and cost-effective filtration systems may help in abating the life-threatening impacts of air pollution.

VI. ACKNOWLEDGEMENTS

VII. REFERENCES

[1] World Health Organization. (2020). *https://www.who.int/health-topics/air-pollution*

[2] Xing, Y., Xu, Y., Shi, M., Lian, Y., "The impact of PM2.5 on the human respiratory system," in *J Thorac Dis*. 2016 Jan; 8(1): E69–E74.

[3] Wu et al., "Air pollution and COVID-19 mortality in the United States: Strengths and limitations of an ecological regression analysis," in *Sci. Adv*. 2020; 6 : eabd4049 4 November 2020.

[4] "Abatement Technologies. Facts about true HEPA filtration", *https://www.abatement.com/learning-center/patient-isolation/facts-about-hepa-filtration/*, 2018.

[5] Brook, D., "What is an air ionizer? Are air ionizers good for you?", *https://breathequality.com/ionizer/* (2019).

[6] Vyas, S., Srivastav, N., Spears, D., "An experiment with Air Purifiers in Delhi during Winter 2015-2016", *PLoS ONE* 11(12): e0167999. doi: 10.1371/journal.pone.0167999

[7] Xia, X., Riviere, N., Mathur, S., Song, X., Xiao, L., Oldenberg, S., Fadeel, B., Riviere, J., "Mapping the surface adsorption forces of nanomaterials in biological systems", *National Institute of Health,* 5(11), 9074-9081, 2011. doi: 10.1021/nn203303c

[8] Zhong, Z., Xu, Z., Sheng, T., Yao, J., Xing, W., Wang, Y., "Unusual Air Filters with Ultrahigh Efficiency and Antibacterial Functionality Enabled by ZnO Nanorods", *ACS Applied Materials & Interfaces*, 7 (38), 21538-21544. doi:10.1021/acsami.5b06810

[9] Wongwatcharapaiboon, J., Gan, G., Riffat, S., "A new air PM2.5 filtrative lamp with a combination of fabric filter and TiO2 coating mop", *International Journal of Low-Carbon Technologies*, 14 (3), 394–399, https://doi.org/10.1093/ijlct/ctz027, 2019.

[10] Ruan, D., Qin, L., Chen, R. et al., "Transparent PAN:TiO2 and PAN-co-PMA:TiO2 Nanofiber Composite Membranes with High Efficiency in Particulate Matter Pollutants Filtration", *Nanoscale Res Lett.*, 15, 7. doi: 10.1186/s11671-019-3225-2.

[11] Szczęśniak, B., Choma, J., & Jaroniec, M., "Gas adsorption properties of graphene-based materials", Advances in Colloid and Interface Science, 243(1), 46-59, 2017. doi: 10.1016/j.cis.2017.03.007

[12] Zhang, S., Sun, J., Hu, D., Xiao, C., Zhuo, Q., Wang, J., Qin, C., Dai, L., "Large-sized graphene oxide/modified tourmaline nanoparticle aerogel with stable honeycomb-like structure for high-efficiency PM2.5 capture", *J. Mater. Chem. A*, 2018, 6, 16139-16148. doi: 10.1021/acsami.8b22382

[13] Zhao. D., Parham. A., Stephens. B. (2015). Evaluating the long-term health and economic impacts of central residential air filtration for reducing premature mortality associated with indoor fine particulate matter (PM2.5) of outdoor origin. *Int. J. Environ. Res. Public Health*, 12(7), 8448-8479.

[14] Giovanetti et al., "Recent advances in graphene based TiO2 nanocomposites (GTiO2Ns) for photocatalytic degradation of synthetic dyes", *Catalysts*, 7(10), 305, 2017. doi:10.3390/catal7100305

[15] Weiss, C. et al. (2020). Toward Nanotechnology-Enabled Approaches against the COVID-19 Pandemic, *ACS Nano* 2020 14 (6), 6383-6406, doi: 10.1021/acsnano.0c03697

[16] Mitra, P. et. al. (2018). Antibacterial and Photocatalytic Properties of ZnO–9-Aminoacridine Hydrochloride Hydrate Drug Nanoconjugates. *ACS Omega* 2018 3 (7), 7962-7970, doi: 10.1021/acsomega.8b00568

[17] Xiong, H.-M. (2013). ZnO Nanoparticles Applied to Bioimaging and Drug Delivery. *Adv. Mater*. 25, 5329-5335. doi: 10.1002/adma.201301732

Pecan: A Novel Approach to Energy Supply and Demand Forecasting in a Photovoltaic Microgrid

Andy Xu

Abstract— **The adoption of renewable energy is crucial to curbing carbon emissions. Localized on-site generation methods such as microgrids are implemented due to their improved reliability and the ease of inclusion of renewable energy generation. However, current forms of renewable energy generation are unreliable. Accurate forecasting of both energy supply and demand are crucial in the transition towards a renewable energy grid and reducing reliance on fossil fuel reserves. Pecan is a novel solution that combines custom deep learning models for energy supply and demand forecasting with an artificial neural network solution. Pecan uses a novel loss function to prioritize grid stability while simultaneously decreasing carbon emissions through a lower error rate. Mean absolute percentage error (MAPE) was used to measure model performance and calculate emission reductions. The supply forecasting prediction from Pecan achieved a MAPE of 1.17%, and the demand forecasting prediction achieved a MAPE of 1.05%. The improved performance of Pecan increases the feasibility and profitability of microgrids and renewable energy solutions.**

I. INTRODUCTION

Due to growing concerns regarding the effects of fossil fuel emissions, the importance of renewable energy resources has grown markedly in recent years. Currently, the industry standard for power management is a centralized power grid largely dependent on fossil fuels. However, centralized power grids present two major problems. First, the rigid, inflexible centralized grid is unable to accommodate the unpredictable nature of current distributed energy resources(DERs). Second, energy is often lost when travelling large distances between energy generation and consumption locations [1]. Distributed, or on site generation, has been proposed as a next generation smart grid solution. This method proposes advantages due to its ability to generate energy locally, greatly reducing the energy lost in transmission, and its superior reliability and resilience due to its small scale and isolation [2].

Short-term energy supply and demand forecasting are necessary to make informed and reliable decisions for distributed energy systems [3]. Currently, reserve scheduling ensures that there are adequate reserves in place when energy demand exceeds supply. These reserves are most commonly fossil fuels due to the necessity of immediate generation. Improvements in energy supply and demand forecasting have the potential to greatly reduce the necessity of fossil fuel reserves and fossil fuel energy generation as a whole. They also have the potential to increase grid stability by greatly reducing potential mismatches of supply and demand.

Multiple deep learning techniques have been proposed in the past for energy supply and demand forecasting. These include artificial neural networks (ANN), convolutional neural networks (CNN), recurrent neural networks (RNN), Long short-term memory networks (LSTM) and bidirectional long short-term memory networks (BLSTM) [4].

The purpose of Pecan is to develop a comprehensive deep learning solution for energy supply and demand in a microgrid that is more accurate and reliable than industry standards. In the present study, a novel method is developed to connect energy supply and demand forecasts to produce an intelligent demand response.

II. METHODS

The primary microgrid dataset used for both energy supply and demand predictions was the UC San Diego Microgrid [5]. For this dataset, 48 hours of energy demand data in kilowatts was inputted into the energy model. The month, day, and hour were used as inputs due to strong seasonal correlations between time and energy consumption. Temperature, taken from the National Oceanic and Atmospheric Administration(NOAA) located at the San Diego Airport, was also used as an input.

The energy generation forecasting model took in the energy generation from distributed solar PV generators throughout the San Diego Microgrid. The model inputted multiple solar irradiance metrics, including Global Horizontal Irradiance (GHI), Direct Normal Irradiance (DNI), and Diffuse Horizontal Irradiance (DHI). The same temporal variables as the energy load forecasting model were used for the energy generation forecasting model.

Both energy supply and demand forecasting models used the same novel deep learning model structure to achieve a lower forecasting error. The model contains a LSTM layer, followed by 3 bidirectional LSTM layers, followed by another LSTM layer, then finally followed by 2 artificial neural network layers. Dropout and batch normalization was used between each layer to prevent overfitting of the model.

The data was split into two parts with 90% being used for training the model and the remaining 10% being used for validation and testing the model accuracy. Temporal isolation was ensured to prevent overfitting and combat bias. To measure model performance, mean absolute percentage error (MAPE) was used. Pecan was compared to a standard ANN

model, a smart persistence model, and the current industry standard. The smart persistence model used the data from the last 24 hours to generate a prediction of equal value for the next 24 hours.

III. RESULTS

The energy generation forecasting portion of Pecan achieved a Mean Absolute Percentage Error (MAPE) of 1.13%. This was an 87% decrease compared to the Smart Persistence model, a 75% decrease in MAPE compared to the Artificial Neural Network (ANN) model and a 62% decrease in MAPE compared to the industry standard models.
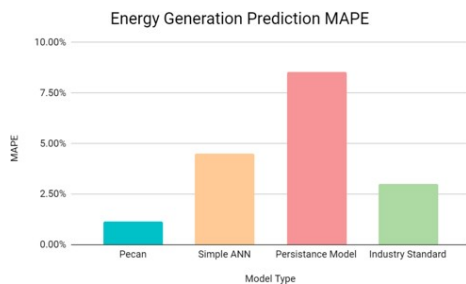

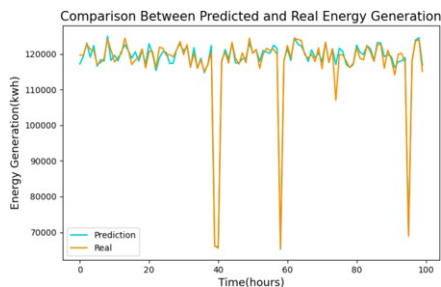Figure 1.   Energy Generation Prediction Model Comparison

Figure 2.   Energy Generation Prediction and Real Value Comparison Visualized

The load forecasting portion of Pecan achieved a MAPE of 1.05%. This was an 89% decrease compared to the Smart Persistence model, a 72% decrease in MAPE compared to the Artificial Neural Network(ANN) model and a 65% decrease in MAPE compared to the industry standard models.
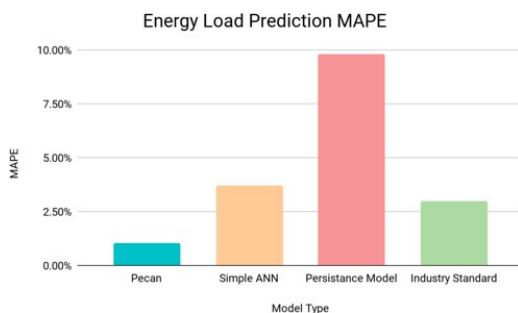

Figure 3.   Energy Load Prediction Comparison
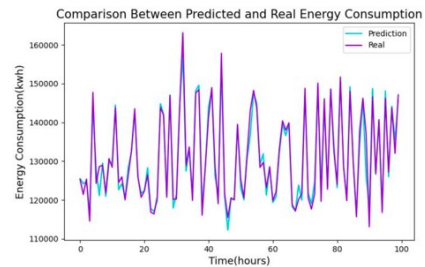
Figure 4.   Energy Load Prediction and Real Value Comparison Visualized

IV. DISCUSSION

The development of Pecan demonstrates the applicability of a novel neural network solution to energy supply and demand predictions. Pecan's more accurate supply and demand forecasts can be used to improve accuracy in the unit commitment problem, thereby increasing forecasting reliability, and greatly improving grid stability and national security. Pecan also has large-scale economic benefits, as improvements in energy forecasting accuracy allow for improved energy price forecasting, as well as lower energy purchasing from the central grid. Pecan helps to increase the feasibility and profitability of microgrids and renewable energy. Future research can build off of Pecan's novel network structure to test on a wider range of microgrid datasets to ensure generalizability, and can apply Pecan's novel structure to other forms of distributed energy resources

REFERENCES

[1]   J. Silvente, G. M. Kopanos, E. N. Pistikopoulos, and A. Espuña, "A rolling horizon optimization framework for the simultaneous energy supply and demand planning in microgrids," *Applied Energy*, vol. 155, pp. 485–501, Oct. 2015, doi: 10.1016/j.apenergy.2015.05.090.

[2]   E. Mengelkamp, J. Gärttner, K. Rock, S. Kessler, L. Orsini, and C. Weinhardt, "Designing microgrid energy markets: A case study: The Brooklyn Microgrid," *Applied Energy*, vol. 210, pp. 870–880, Jan. 2018, doi: 10.1016/j.apenergy.2017.06.054.

[3]   F. Ünal, A. Almalaq, and S. Ekici, "A Novel Load Forecasting Approach Based on Smart Meter Data Using Advance Preprocessing and Hybrid Deep Learning," *Applied Sciences*, vol. 11, no. 6, p. 2742, Jan. 2021, doi: 10.3390/app11062742.

[4]   S. Aslam, H. Herodotou, S. M. Mohsin, N. Javaid, N. Ashraf, and S. Aslam, "A survey on deep learning methods for power load and renewable energy forecasting in smart microgrids," *Renewable and Sustainable Energy Reviews*, vol. 144, p. 110992, Jul. 2021, doi: 10.1016/j.rser.2021.110992.

[5]   S. Silwal, C. Mullican, Y.-A. Chen, A. Ghosh, J. Dilliott, and J. Kleissl, "Open-source multi-year power generation, consumption, and storage data in a microgrid," *Journal of Renewable and Sustainable Energy*, vol. 13, no. 2, p. 025301, Mar. 2021, doi: 10.1063/5.0038650.

Modeling the Stellar Kinematics of the Thick Disk and Halo of the Andromeda Galaxy

Yaqoub Ahmad, Jessamine Qu, and Penelope Strong

Abstract— **In studying the Andromeda Galaxy to better grasp its physical components—particularly its northeastern region—we utilized Python code to simulate its halo and disk. Using previously observed data and various formulas, such as one that calculates the density distribution of stars, we closely modeled the real dimensions of Andromeda's halo and disk. Moreover, looking at different velocity dispersions along the height, radius, and angle axis helped us further understand Andromeda's actual dispersions. Comparing models with different percentages of stars in the thick disk and halo, we found that both the thick disk and the halo had minimal effect on the observed dispersion. Furthermore, we had difficulty observing overarching dispersion trends brought about by changing the dispersion coordinate variables. Based on these observations, a more natural and substantial dispersion of Andromeda can be concluded. As we continue, our analysis will assist us in fine-tuning the model, more accurately simulating the Andromeda Galaxy, and eventually adapting the code to forward model any galaxy.**

I. INTRODUCTION

As the brightest and largest galaxy within our own cluster of galaxies, the Andromeda Galaxy has its own unique attributes [1]. Yet, this huge collection of gas, dust, and billions of stars, along with their solar systems, is identical to every other galaxy in that it is held together by gravity. Andromeda has three main parts—the bulge, a dense ball of stars at its center; the disk, a thick and less dense disk of stars; the halo, a dimmer, sparsely populated cluster of stars surrounding it. Because Andromeda is nearby and frequently observed, it is often used as a starting point for galaxy models [4]. Specifically, previous research conducted by J. Veljanoski, among others, examined the velocity dispersion of Andromeda's outer halo globular cluster system and discusses various formation scenarios for Andromeda [3]. Therefore, there is a considerable amount of information available regarding Androme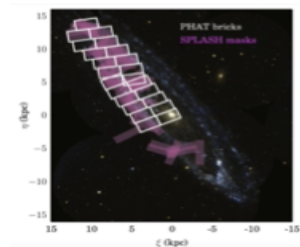da. However, we remain unsure about much of the details of its evolution. Understanding the dynamics of its stellar population could help us understand the history of Andromeda's formation, as well as the formation of other large galaxies, including our own. Likewise, there remains debate over the external and internal factors begetting dynamical heating of stellar disks, including the central bar, dwarf satellite bombardment, molecular clouds, and dark matter substructure. Hence, we are

Figure 1. White outline indicates the observed (northeastern) region of Andromeda. [2]

coding a model illustrating the stellar kinematics of the thick disk and halo of the Andromeda galaxy.

II. METHODS

By creating a simulation of Andromeda via Python code, we will eventually be able to forward model any galaxy, helping us better understand far off galaxies as well as our very own Milky Way. We fine-tuned the code of our model to Andromeda in particular because of its close proximity to us, similarity to the Milky Way, and familiarity to astronomers. Our model was created with Jupyter Notebook and implements the libraries NumPy, to perform operations on arrays, and Matplotlib.pyplot, to generate graphs of the galaxy. However, this model only includes the disk and halo. We also limited the code to Red Giant Branch Stars within the Northeastern part of Andromeda (Fig. 1) as they are larger, brighter, and therefore easier to observe than stars like our sun. Hence, we modeled the dispersion of velocities of Red Giant Branch Stars in Andromeda's Northeastern disk and halo. Both the simulated halo and disk are projected in cylindrical units. However, the spherical halo is modeled in two dimensions because it appears as a circle no matter what angle it is viewed at, whereas the disk is projected in three-dimensional space due to its thickness. To allow for direct comparison between the physical quantities generated by our simulation and the spectral line profiles shown by observations, we must create a model to generate data that can be closely compared with observed data. This technique is known as forward modeling [6]. Utilizing this approach, we built our code around known parameters and then adjusted the unknown parameters to match observed data. Initially, we had Andromeda's sky position and two velocity components as known parameters. These variables were included in existing code that modeled the distribution of Red Giant Branch Stars and the dispersion of their radial velocities in Andromeda's disk. But, the initial model only had one overall dispersion variable and did not include Andromeda's halo.

After subdividing the original dispersion variable into three different variables for dispersion along the phi angle ($!_{..}$), the z coordinate or height ($!_\$$), and the radius ($!_\%$), we used the equation seen in Equation 1: *C* being the expected total dispersion of Andromeda when observed. We input the variable we knew, $!_\$$; plugged in various possible numbers for one unknown, $!_{..}$; and solved for the other, $!_\%$. This resulted in various plausible data sets that we could plug into the model and then compare to Andromeda to help narrow down what Andromeda's actual $!_{..}$ and $!_\%$ are.

$$!_\%^\& + !_{..}^\& + !_\$^\& =)^\&$$

Equation 1. The formula for *C* or the expected total dispersion of Andromeda observed, where σ_ϕ describes dispersion along the phi angle, σ_R describes dispersion along the radius, and σ_z describes dispersion along the z coordinate or height.

We then modeled the halo in two-dimensional space. Equation 2 illustrates the power-law formula we used for determining the density of stars in relation to the radius. This formula accounts for how stars from the halo begin to dominate as the density of the stars from the disk falls off, as indicated from analysis of Andromeda's luminosity profile [4]. Thus, we simulated the stars in the halo according to the distribution defined by Equation 2. On the other hand, the velocity dispersion of the stars in the halo was generated randomly within the bounds of our observed data.

$$R' = R_c \sqrt{(1-w)^{\left(\frac{1}{1+\alpha}\right)} - 1}$$

Equation 2. The formula for the density of stars in relation to the radius, where *R'* is the observed projected radius, *Rc* is the halo core radius, *α* is the halo surface brightness profile power-law index, and *w* represents a random number uniformly distributed between zero and one.

Finally, we periodically adjusted the free parameters and compared the different simulations of Andromeda with the Andromeda observed in the sky. We primarily focused on how stars within the halo impacted the observed dispersion profile. This meant thoroughly analyzing the ratio of halo stars to disk stars and its effect on the simulated galaxy's velocity dispersion. Additionally, we looked for significant trends in dispersion as we altered the σ_R and σ_ϕ variables.

III. RESULTS AND DISCUSSION

Overall, we discovered that the projection effects from the thickness of the disk were less pronounced than we previously thought. This suggested that there is an intrinsic dispersion unaffected by the halo. From these observations, the halo appears to be subdominant, having little to no effect on the dispersion within the surveyed region. Moreover, we found it extremely difficult to perceive overarching trends as σ_ϕ and σ_R changed. This implies that there is no observable dispersion contribution from the thick disk. However, based on our observations, the thick disk and halo do not account for much of Andromeda's overall dispersion. Ultimately, our findings were limited to what we could clearly discern with the human eye. Accurately and precisely observing such large amounts of data within a model displayed on a computer monitor would unavoidably prove difficult. However, unnoticed does not denote unnoticeable. In fact, existent and important trends could be uncovered with further research. By forward-modeling our observations of Andromeda, we are able to better understand the structural and kinematical properties of its thick disk and RGB stars, thereby helping to create a more realistic model which may be applicable to kinematical data of similar, nearby galaxies, such as the

Triangulum galaxy. Additionally, future investigations could focus on determining the effect of Andromeda's bulge component on the galaxy's kinematics.
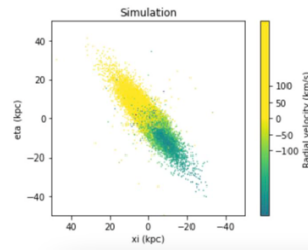

Figure 2. Model of the disk and halo of the Andromeda galaxy with 95% contribution of stars from the disk and 5% from the halo. The xi and eta axis labels are placeholders for coordinates.
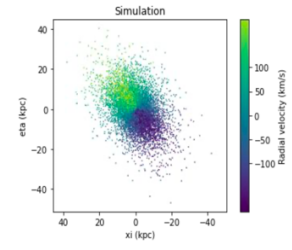
Figure 3. Illustrates the radial velocity within Andromeda's disk where σ_R equals 60, σ_ϕ equals 60.03, and σ_z equals 20. The xi and eta axis labels are placeholders for coordinates.

Our team has been awarded the most time of the Hubble Telescope's upcoming cycle 29, which will allow us to observe a broader area of Andromeda in greater detail. We will continue improving the model, broadening the parameters within Andromeda, and exploring the relationship between σ_R and σ_ϕ.

REFERENCES

[1] Alan W. McConnachie *et al* 2018 *ApJ* 868 55

[2] Kinematics of M31's Thick Disk of Red Giant Stars. P. Guhathakurta, A. Quirk, M. Fardal, K. McConnell, J. Salinas, and C. Villanueva. BAAS, 237, 351.05, 2021.

[3] J. Veljanoski, A. D. Mackey, A. M. N. Ferguson, A. P. Huxor, P. Côté, M. J. Irwin, N. R. Tanvir, J. Peñarrubia, E. J. Bernard, M. Fardal, N. F. Martin, A. McConnachie, G. F. Lewis, S. C. Chapman, R. A. Ibata, A. Babul, The outer halo globular cluster system of M31 – II. Kinematics, *Monthly Notices of the Royal Astronomical Society*, Volume 442, Issue 4, 21 August 2014, Pages 2929–2950, https://doi.org/10.1093/mnras/stu1055

[4] R. Kipper, P. Tenjes, O. Tihhonova, A. Tamm, E. Tempel, Stellar kinematics using a third integral of motion: method and application on the Andromeda galaxy, *Monthly Notices of the Royal Astronomical Society*, Volume 460, Issue 3, 11 August 2016, Pages 2720–2730, https://doi.org/10.1093/mnras/stw1194

[5] Stéphane Courteau *et al*. (2011). The Luminosity Profile and Structural Parameters of the Andromeda Galaxy. *The Astrophysical Journal*, 739. Retrieved from https://iopscience.iop.org/article/10.1088/0004-637X/739/1/20/pdf.

[6] Van Doorsselaere T, Antolin P, Yuan D, Reznikova V and Magyar N (2016) Forward Modeling of EUV and Gyrosynchrotron Emission from Coronal Plasmas with FoMo. *Front. Astron. Space Sci.* 3:4. doi: 10.3389/fspas.2016.00004

Alzheimer's Disease and A Prion-Like Protein: A Toxic Relationship

Krithika Karthik

Abstract— **The role of TDP-43 and its prion-like features in Alzheimer's disease (AD) represents a new avenue of research concerning the pathogenesis of the disorder. Research has focused on identifying proteins involved in inducing aggregation/toxicity of the illness, with the Tau and ß-amyloid proteins being primarily responsible. The TDP-43 protein was first discovered in 1995 and has attracted considerable interest in recent years. This review details the structural and functional characteristics of TDP-43. Special emphasis is given to the post-translational modifications and mutations that accompany neurotoxicity and protein aggregates found in the brain tissue of AD patients. The interface of TDP-43 with other proteins involved in AD progression is also elucidated based on studies in this regard. Investigations using animal models with the intent to identify potential therapeutic strategies to combat the disease have also been outlined in this work.**

I. INTRODUCTION

Alzheimer's disease (AD) is a progressive neurodegenerative disease identified by the death of brain cells, initially observed in the frontotemporal lobes of the brain, characterized by symptoms such as impaired neuronal transmission, brain atrophy and consequent shrinkage [1], [2]. External manifestations of AD include dementia, involving memory loss and declining cognitive and social skills. Of the 50 million cases of dementia worldwide, 60-70% have been diagnosed as Alzheimer's, as indicated by WHO statistics [2].

In 1906, Alois Alzheimer, a clinical psychiatrist and neuropathologist at Frankfurt Psychiatric Hospital, provided the first description of Alzheimer's disease as a 'peculiar severe disease process of the cerebral cortex' [3], and this disease continues to remain a mystery in some major aspects. Brain scans of AD patients have revealed extracellular parenchymal and intraneuronal aggregates of proteins, primarily the beta-amyloid and Tau proteins, leading to the formation of amyloid-beta (Aß) plaques and neurofibrillary tangles (NFTs) respectively [4]. Authors have also hypothesized that the onset of the disease could be related to prions, or prion-like polymorphisms of proteins, as observed in the case of the Tau, Aß, or even the TDP-43 protein [5], [6].

Prions are misfolded proteins which have been identified as causative agents of disorders such as Creutzfeldt-Jakob disease. Also, prions multiply via the conformational conversion of normal cellular prion proteins (PrPc) to the disease-causing (PrPSc) isoforms [7], [8] rather than the conventional nucleic acid replication.

Aberrant processing during polypeptide synthesis due to mutations in the prion protein gene dictate the specific abnormality of the neurons in the disorder [9]. In the context of AD, versions of Tau and Aβ proteins are observed to adopt prion-like properties, causing them to spread through the brain to induce neurotoxicity. Thus, authors have suggested that these proteins could be potential culprits for progression of the disease [10].

II. EFFECT OF MUTANT PROTEINS

Familial Alzheimer's disease (FAD) is an autosomal dominant disease. Mutations in the amyloid precursor protein (APP), presenilin 1 or 2 genes have been reported to result in FAD [9]. APP undergoes hydrolysis to form Aβ peptides that are responsible for the amyloid fibrils found in the plaques of AD brains [11]. APP is cleaved at residue 671 by β-secretase to form the Aβ (1-40) fragment and at residue 711 or 713 by γ-secretase to form the Aβ (1-42) fragment as depicted in Fig. 1 [9]. These Aβ (1-42) peptides can cause disruption in the central nervous system [9], [12] whereas the Aβ (1-40) fragment is mainly present in the amyloid fibrils of the plasma and cerebrospinal fluid [12]. Plaque formation is reported due to Aβ peptides that are influenced by specific mutations (APP695, APP751, APP770) of the truncated APP gene, which are commonly located at cleavage sites of β-secretase and γ-secretase [13].
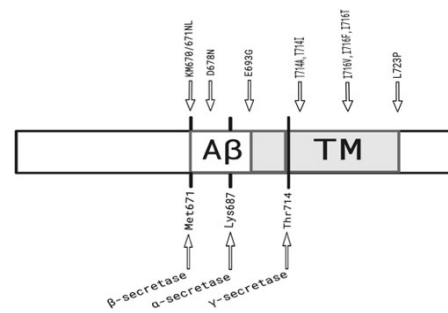
Figure 1. The C-terminus of the APP protein. The secretase cleavage sites and some mutations in the protein are depicted. Aβ, amyloid-beta; TM, transmembrane domain (Reproduced from [13])

Apart from the two key proteins Tau and Aβ, TDP-43 has also been implicated in the neurodegenerative processes occurring in AD [11]. For these reasons, studies conducted on TDP-43 have been reviewed in order to explore the likelihood of this protein as a potential therapeutic target.

III. TDP-43 PROTEIN

Structure

The TAR DNA binding protein (TDP-43) of 43 kDa is encoded by the TARDBP gene of 6 exons at the 1p36.22 locus [14]. This heterogenous ribonucleoprotein (hnRNP) of 414 amino acids [15] is localized primarily in the nucleus [16]. It consists of an N-terminal domain (NTD) and C-terminal domain (CTD), which is characterized as a prion-like domain [16]. The protein also has 2 RNA recognition motifs (RRM), and a nuclear localization signal (NLS) domain. The structure is depicted in Fig. 2.
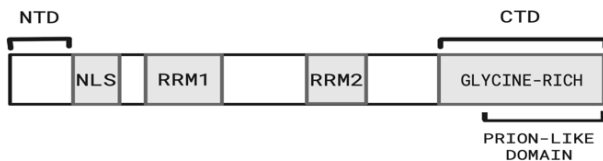
Figure 2. The TDP-43 protein structure and important domains
(Reproduced from [17])

Function

The overall function of TDP-43 under nontoxic conditions includes the RNA regulation during transcription, mRNA stabilization, and alternative splicing [16]. The specific function of each domain is detailed in Table 1. The CTD is a low complexity sequence that can transition between alpha-helix and beta-sheet structures, similar to how prions trigger the onset of a disease [1], [18]. This, along with the prevalence of pathogenic TDP-43 in 20-50% of AD cases [11], has led researchers to suspect that TDP-43 could also be a possible contributor to the toxicity observed in AD brains [16].

TABLE I. STRUCTURAL AND FUNCTIONAL ANALYSIS OF THE
TDP-43 PROTEIN DOMAINS [19]

| Domain	Residual Position	Function under Non-toxic Conditions
NTD	1-77	Protein dimerization and oligomerization [16], [20]
NLS	78-100	Translocation of TDP-43 from nucleus to cytoplasm for cytoplasmic accumulation of proteins [4], [15]
RRM1	106-177	Specific RNA binding [4], [15]
RRM2	192-259	Specific RNA binding [4], [15]
CTD	260-414	Phase separation, aggregation, solubility, and protein homeostasis [16], [20]

Proteinopathies involving TDP-43

Studies have shown strong evidence for TDP-43 as a pathological hallmark for frontotemporal lobar degeneration (FTLD) and amyotrophic lateral sclerosis (ALS) [21]. TDP-43 localization in AD brains has also received attention in recent times. TDP-43 deposits as inclusion bodies in neuronal and glial cells of the central nervous system and spreads through the brain in 6 stages. The accumulation begins in the amygdala (stage 1), followed by the subiculum entorhinal cortex (stage 2), the hippocampus and occipitotemporal cortex (stage 3), the ventral striatum, insular and temporal cortices (stage 4), the brainstem (stage 5) and finally the basal ganglia and midfrontal cortex (stage 6) as reported by [6], [11].

The TARDBP gene was found to have mutations in genetic cases of ALS and FTLD [16], [22], but few researchers have assessed this gene in relation to AD [1]. A clinical trial consisting 181 AD patients and 130 controls in a Japanese population with 8 TARDBP gene polymorphisms observed no significant relationship [1]. Yet, it is still likely that the TDP-43 and Aβ/tau protein are related. Results show that Aβ deposits cause can cause abnormal TDP-43 aggregation and that TDP-43 is involved in tau aggregation [11]. Thus, TARDBP gene mutations could be responsible for the abnormal, mutated behavior of TDP-43, indicative of the need for more research on TDP-43 mutations pertaining to AD.

III. POST-TRANSLATIONAL MODIFICATIONS (PTMs) OF TDP-43

The TDP-43 protein can undergo multiple PTMs including truncation, ubiquitination, phosphorylation, and acetylation. As a target for multiple PTMs TDP-43 may contribute to AD progression in multiple ways. It should be noted that although there is evidence suggesting TDP-43 PTMs can stimulate aggregation [19], some studies also report PTMs occurring post-aggregation (refer to the Phosphorylation section) [23]. Consequently, whether these PTMs are the cause or the effect of aggregations is still unclear.

Truncation

The cleavage of TDP-43 leads to the formation of N terminal fragments (NTFs) or C terminal fragments (CTFs), depending on the cleavage site. The NTFs retain their function and break down in the nucleus, whereas CTFs translocate to the cytoplasm and aggregate [24]. The truncation of the RRMs affects the ability of the protein to carry out RNA regulation/binding and dimerization. Furthermore, the removal of part of/the whole NLS domain triggers further cytoplasmic protein aggregation [19]. Proteolytic cleavage at the 89/90 position and at position 169/170 or 174/175 produces TDP-35 and TDP-25, which are other commonly occurring fragments [19]. As studied by Li et al., western blot analyses of such TDP-43 fragments show higher levels of insolubility and therefore propensity to aggregate [23].

Ubiquitination

Ubiquitination involves the binding of the ubiquitin protein to TDP-43 [19]. This PTM regulates the activation/inactivation, localization and interactions of proteins. Common ubiquitination sites are lysine residues,

particularly at positions 160, 181, and 263. François-Moutal et al. found that the K181 residue could disrupt the secondary/tertiary structure of TDP-43 due to the changes in the interactions between the RRM domains or between the RRMs and NTD. It is also observed that the K263 residue can be easily ubiquitinated to decrease levels of RNA binding and aid in aggregative tendencies [19].

Phosphorylation

Phosphorylation involves the addition of phosphate groups to amino acids in the TDP-43 sequence. Abnormal or hyper-phosphorylation is also known to indicate neurotoxicity [19]. In fact, hyperphosphorylation of the tau protein in AD triggers the formation of NFTs [11], [25]. Serine residues, and less commonly threonine residues, are common phosphorylation sites. Interestingly, the CTD contains such serine and threonine residues [18]. This suggests that TDP-43 phosphorylation promotes aggregation and neurodegeneration [24]. However, the effect of this PTM is debatable, with certain studies suggesting the PTM serves a cytoprotective function, easing proteotoxic stress caused by accumulation of misfolded proteins and aggregates [24]. Either way, phosphorylation appears to play a key role in the progression of TDP-43 proteinopathies in a disease like AD.

Acetylation

Acetylation involves the addition of acetyl groups to amino acids in the TDP-43 sequence. The common target for acetylation are lysine residues [19]. This process can occur via various mechanisms such as RNA processing, cytoskeleton association, and cellular signaling [19]. The specific position undergoing acetylation could be Lys82 and 192 [19]. Cohen et al. have reported the occurrence of inclusion bodies in the cytoplasm reminiscent of TDP-43 proteinopathy aggregates [21].

IV. STUDIES INVOLVING ANIMAL MODELS OVEREXPRESSING PROTEINS IMPLICATED IN AD

Transgenic Mice Models: Advantages and Limitations

Mice are commonly used animals in studies due to their genetic similarity to humans. Wild type mice were found to exhibit 97% and 88% sequence homology to the human version of the APP and Tau proteins respectively [4]. When these mice overexpress APP, their pathology is reminiscent of AD pathogenesis. The Aß plaques are composed of the Aß (42) peptide, ubiquitin, and α-synuclein, among other constituents. α-synuclein is of special note as it has been characterized as "prion-like" following studies using animal models [12]. The mice, like human patients, were also observed to develop cognitive impairments which were not directly proportional to Aß plaque formation [26].

However, although mice can produce Aß, the plaque formation and development of AD-like characteristics are dependent on the APP mutations which resemble FAD, but not sporadic AD cases. The development of NFTs is also absent in these models. Nonetheless, if mutant APP and several other proteins are present, age-dependent development of both Aß plaques and NFTs can form in mice [26], [4].

Non-human Primate Models: Advantages and Limitations

Nonhuman primates are preferable models to study disease development in that their "behavioral complexity," brain size, and genetics are highly similar to that of humans. In fact, their Aß and tau protein have 100% and 99.5-100% sequence homology with the human forms respectively, and protein accumulation occurs naturally. New world monkeys like squirrel monkeys experience neurotoxicity resembling that of humans, with possession of Aß (1-40) and Aß (1-42) of particular interest. These peptides accumulate and aggregate to form Aß plaques [4]. Thus, these models are favorable to study the propagation of the Aß protein and plaque formation.

However, primates suffer from the limitation of long lifespans and delayed neuropathology. Furthermore, cognitive symptoms are not evident and Aß and tau are found in smaller amounts than as seen in humans. Baboons, for example, only rarely have Aß accumulation and squirrel monkeys have no NFTs present [4].

Few studies have been done on TDP-43 using mammalian models. Mutant TDP-43 isoforms mainly remain in the nucleus or thinly distributed near the nucleus/in the cytoplasm in mice, but shuttle to the cytoplasm in primates [27], [24]. This distinction in the distribution of TDP-43 between mice and monkeys is to be noted, as it calls the reliability of the models into question.

V. THERAPEUTIC APPROACHES TARGETING PROTEINS IMPLICATED IN AD

Drugs can be used to target various aspects of AD, including preventing APP or TDP-43 fragmentation, reducing TDP-43 expression, or inhibiting PTMs of tau and TDP-43. In this review, the drugs targeting various aspects of TDP-43 proteinopathies are detailed in Table 2 [15].

TABLE II. DRUGS TARGETING TDP-43

Drug	Description	Model System Used	Effects Observed (Interactions with TDP-43)
Berberine	Medicinal herb that can be orally ingested [28]	Cell culture model Mouse model	Reduction of accumulation and aggregation of TDP-43 fragments [28] Decrease in levels of Aß and phosphorylated tau, leading to improvements in cognitive symptoms (learning/motor skills and spatial

			memory) [29], [28]
N-Acetyl-cysteine (NAC)	Compound with antioxidative properties [15]	-	Modification of abnormal cytoplasmic accumulation of TDP-43 in neuron-like cells and reduction of toxic effects due to arsenite-induced insolubility and ubiquitination [15]
Dexamethasone	Synthetically produced steroid hormone (glucocorticoid) [15]	Transgenic mice model	Increase in TDP-25 solubility, improving cognitive symptoms [15]
Epigallocatechin gallate (EGCG)	Polyphenolic plant compound found [30]	-	Conversion of Aß protein, synuclein protein (eg- α-synuclein) and yeast prion (PSI), which tend to form amyloid deposits, into harmless oligomers [15]
Curcumin Dimethoxy	A compound, specifically monocarbonyl dimethoxy [15]	-	Reduce toxicity induced by mutant TDP-43 and halt aggregation due to pathological TDP-25 [15]
QBP1 (PolyQ peptide binding 1)	Peptide sequence targeting polyglutamine sequences reminiscent of amyloid fibrils [31]	*In vitro* models	Binding and inhibition of amyloid fibril production by targeting the Q/N rich CTD of TDP-43 [15]
Rolipram	Drug capable of inhibiting phosphodiesterase-4 enzyme	Transgenic mice models	Decrease in levels and aggregation of TDP-43 in specific neurons [15]
Riluzole	Drug clinically approved to treat ALS [15]	-	Decrease in TDP-43 interactions in a dose-dependent fashion [32]

VI. DISCUSSION

Studies investigating the role of TDP-43 in AD is still relatively minimal. Research has mainly focused upon ALS and FTLD, with relatively few studies directly extending this research to TDP-43 proteinopathies occurring in AD. Further studies are required to determine any reliable correlations between AD and TDP-43. The regulation of the Aβ/tau-protein by TDP-43 is unclear, so the interactions of TDP-43 requires further exploration [11]. This different perspective,

of investigating the prion-like properties of the TDP-43 protein, should be seriously accounted for as it could help hasten the development of safer, more effective treatment methods for AD.

Immunotherapy appears to be a promising path [9], [12]. For example, in one study [9], [26], the administration of antibodies decreased Aβ plaque levels and improved the cognitive symptoms. However, any immunotherapeutic treatment comes with a high risk factor. One example of a problem observed is vasogenic edema or encephalitis, which in some cases is also accompanied by hemorrhage [4]. Some drugs are promising, however, especially those that can slow the 6-stage transgression of TDP-43 through the brain. Such a drug could have important implications as the majority of AD diagnoses occur late into the disease—80% in stage 3 and 85% in stage 4 or 5 [6]—and slowing movement of TDP-43 can allow for earlier detection of proteinopathies. It should also be noted that certain drug molecules described in this review, such as NAC, are nutritional supplements and their potential as preventative/protective drugs should be explored in greater detail.

Caenorhabditis elegans models have been used extensively to study AD, contributing to major advances in the repertoire of this disease [22], [33]. *Drosophila melanogaster* is another common model organism. The *C. elegans* worms and *D. melanogaster* flies can be rapidly multiplied at low costs [33], [34], such that they are a convenient choice for *in vivo* studies, although phenotypic features may not resemble that of humans. For mammalian models, numerous ethical guidelines must be considered. Therefore, there is no ideal model organism to study AD, so the model should be selected cognizant of what aspect needs to be studied.

Mutations and PTMs play vital roles on the behavior of TDP-43 and its consequent effect on AD cases. Most cases of AD have been associated with mutations in the APP, PSEN1, PSEN2 genes among others. These mutations should be studied further, including their interaction with the TARDBP gene. The interaction of different PTMs could also be explored regarding effects on pathogenesis. as ubiquitination and truncation appear to drive further neurotoxicity together [18]. Overall, the PTMs of TDP-43 seem to be interrelated and codependent in certain situations, and hence they pave the way for the use of interesting study designs in order to further understand the role of TDP-43 in AD.

VII. CONCLUSION

The gravity of this situation, in which both the diagnosis and treatment/prevention of AD is unclear, is not to be understated. Different mechanisms by which drugs can target TDP-43 in AD brain tissues must be explored further. The relationship between TDP-43 and AD is not definite but has strong supporting evidence, indicating that more research is required to directly associate both. Although TDP-43 is not prevalent in all AD cases, it presents a novel approach due to

its prion-like characteristics. The TDP-43 protein could therefore pioneer breakthroughs in studies directed towards a more comprehensive knowledge of AD with the objective to develop possible prevention/therapeutic strategies.

VIII. ACKNOWLEDGEMENT

IX. REFERENCES

[1]   A. C. Wilson, B. N. Dugger, D. W. Dickson, and D.-S. Wang, "TDP-43 in aging and Alzheimer's disease - a review," *Int. J. Clin. Exp. Pathol.*, vol. 4, no. 2, pp. 147–155, Jan. 2011.

[2]   "Dementia." https://www.who.int/news-room/fact-sheets/detail/dementia

[3]   H. Hippius and G. Neundörfer, "The discovery of Alzheimer's disease," *Dialogues Clin. Neurosci.*, vol. 5, no. 1, pp. 101–108, Mar. 2003, doi: 10.31887/DCNS.2003.5.1/hhippius.

[4]   E. Drummond and T. Wisniewski, "Alzheimer's disease: experimental models and reality," *Acta Neuropathol. (Berl.)*, vol. 133, no. 2, pp. 155–175, Feb. 2017, doi: 10.1007/s00401-016-1662-x.

[5]   K. A. Josephs *et al.*, "Staging TDP-43 pathology in Alzheimer's disease," *Acta Neuropathol. (Berl.)*, vol. 127, no. 3, pp. 441–450, Mar. 2014, doi: 10.1007/s00401-013-1211-9.

[6]   S. Nag, L. Yu, P. A. Boyle, S. E. Leurgans, D. A. Bennett, and J. A. Schneider, "TDP-43 pathology in anterior temporal pole cortex in aging and Alzheimer's disease," *Acta Neuropathol. Commun.*, vol. 6, no. 1, p. 33, May 2018, doi: 10.1186/s40478-018-0531-3.

[7]   H. Liu *et al.*, "Solution Structure of Syrian Hamster Prion Protein rPrP(90−231)," *Biochemistry*, vol. 38, no. 17, pp. 5362–5377, Apr. 1999, doi: 10.1021/bi982878x.

[8]   R. Riek, S. Hornemann, G. Wider, M. Billeter, R. Glockshuber, and K. Wüthrich, "NMR structure of the mouse prion protein domain PrP(121–231)," *Nature*, vol. 382, no. 6587, pp. 180–182, Jul. 1996, doi: 10.1038/382180a0.

[9]   S. B. Prusiner, "Neurodegenerative Diseases and Prions," *N. Engl. J. Med.*, vol. 344, no. 20, pp. 1516–1526, May 2001, doi: 10.1056/NEJM200105173442006.

[10]   "Alzheimer's Disease is a 'Double-Prion Disorder,' Study Shows." https://www.ucsf.edu/news/2019/05/414326/alzheimers-disease-double-prion-disorder-study-shows

[11]   W. Huang *et al.*, "TDP-43: From Alzheimer's Disease to Limbic-Predominant Age-Related TDP-43 Encephalopathy," *Front. Mol. Neurosci.*, vol. 13, p. 26, 2020, doi: 10.3389/fnmol.2020.00026.

[12]   F. Ciccocioppo *et al.*, "Neurodegenerative diseases as proteinopathies-driven immune disorders," *Neural Regen. Res.*, vol. 15, no. 5, pp. 850–856, May 2020, doi: 10.4103/1673-5374.268971.

[13]   L. M. Bekris, C.-E. Yu, T. D. Bird, and D. W. Tsuang, "Genetics of Alzheimer disease," *J. Geriatr. Psychiatry Neurol.*, vol. 23, no. 4, pp. 213–227, Dec. 2010, doi: 10.1177/0891988710383571.

[14]   H.-Y. Wang, I.-F. Wang, J. Bose, and C.-K. J. Shen, "Structural diversity and functional implications of the eukaryotic TDP gene family," *Genomics*, vol. 83, no. 1, pp. 130–139, Jan. 2004, doi: 10.1016/S0888-7543(03)00214-3.

[15]   E. Buratti, "Targeting TDP-43 proteinopathy with drugs and drug-like small molecules," *Br. J. Pharmacol.*, vol. 178, no. 6, pp. 1298–1315, Mar. 2021, doi: 10.1111/bph.15148.

[16]   M. Jo, S. Lee, Y.-M. Jeon, S. Kim, Y. Kwon, and H.-J. Kim, "The role of TDP-43 propagation in neurodegenerative diseases: integrating insights from clinical and experimental studies," *Exp. Mol. Med.*, vol. 52, no. 10, pp. 1652–1662, Oct. 2020, doi: 10.1038/s12276-020-00513-7.

[17]   S. Farina, F. Esposito, M. Battistoni, G. Biamonti, and S. Francia, "Post-Translational Modifications Modulate Proteinopathies of TDP-43, FUS and hnRNP-A/B in Amyotrophic Lateral Sclerosis," *Front. Mol. Biosci.*, vol. 8, p. 693325, Jul. 2021, doi: 10.3389/fmolb.2021.693325.

[18]   R. C. Hergesheimer *et al.*, "The debated toxic role of aggregated TDP-43 in amyotrophic lateral sclerosis: a resolution in sight?," *Brain J. Neurol.*, vol. 142, no. 5, pp. 1176–1194, May 2019, doi: 10.1093/brain/awz078.

[19]   L. François-Moutal, S. Perez-Miller, D. D. Scott, V. G. Miranda, N. Mollasalehi, and M. Khanna, "Structural Insights Into TDP-43 and Effects of Post-translational Modifications," *Front. Mol. Neurosci.*, vol. 12, pp. 301–301, Dec. 2019, doi: 10.3389/fnmol.2019.00301.

[20]   B. Borroni, A. Alberici, and E. Buratti, "Review: Molecular pathology of frontotemporal lobar degenerations," *Neuropathol. Appl. Neurobiol.*, vol. 45, no. 1, pp. 41–57, Feb. 2019, doi: 10.1111/nan.12534.

[21]   T. J. Cohen, A. W. Hwang, C. R. Restrepo, C.-X. Yuan, J. Q. Trojanowski, and V. M. Y. Lee, "An acetylation switch controls TDP-43 function and aggregation propensity," *Nat. Commun.*, vol. 6, no. 1, p. 5845, Jan. 2015, doi: 10.1038/ncomms6845.

[22]   N. F. Liachko, C. R. Guthrie, and B. C. Kraemer, "Phosphorylation Promotes Neurotoxicity in a Caenorhabditis elegans Model of TDP-43 Proteinopathy," *J. Neurosci.*, vol. 30, no. 48, pp. 16208–16219, Dec. 2010, doi: 10.1523/JNEUROSCI.2911-10.2010.

[23]   H.-Y. Li, P.-A. Yeh, H.-C. Chiu, C.-Y. Tang, and B. P. Tu, "Hyperphosphorylation as a Defense Mechanism to Reduce TDP-43 Aggregation," *PLOS ONE*, vol. 6, no. 8, p. e23075, Aug. 2011, doi: 10.1371/journal.pone.0023075.

[24]   D. Chhangani, A. Martín-Peña, and D. E. Rincon-Limas, "Molecular, functional, and pathological aspects of TDP-43 fragmentation," *iScience*, vol. 24, no. 5, p. 102459, May 2021, doi: 10.1016/j.isci.2021.102459.

[25]   M. A. DeTure and D. W. Dickson, "The neuropathological diagnosis of Alzheimer's disease," *Mol. Neurodegener.*, vol. 14, no. 1, p. 32, Dec. 2019, doi: 10.1186/s13024-019-0333-5.

[26]   F. M. LaFerla and K. N. Green, "Animal Models of Alzheimer Disease," *Cold Spring Harb. Perspect. Med.*, vol. 2, no. 11, pp. a006320–a006320, Nov. 2012, doi: 10.1101/cshperspect.a006320.

[27]   P. Yin *et al.*, "Caspase-4 mediates cytoplasmic accumulation of TDP-43 in the primate brains," *Acta Neuropathol. (Berl.)*, vol. 137, no. 6, pp. 919–937, Jun. 2019, doi: 10.1007/s00401-019-01979-0.

[28]   C.-F. Chang *et al.*, "Therapeutic effect of berberine on TDP-43-related pathogenesis in FTLD and ALS," *J. Biomed. Sci.*, vol. 23, no. 1, p. 72, Oct. 2016, doi: 10.1186/s12929-016-0290-z.

[29]   S. S. K. Durairajan *et al.*, "Berberine ameliorates β-amyloid pathology, gliosis, and cognitive impairment in an Alzheimer's disease transgenic mouse model," *Neurobiol. Aging*, vol. 33, no. 12, pp. 2903–2919, Dec. 2012, doi: 10.1016/j.neurobiolaging.2012.02.016.

[30]   Q. Y. Eng, P. V. Thanikachalam, and S. Ramamurthy, "Molecular understanding of Epigallocatechin gallate (EGCG) in cardiovascular and metabolic diseases," *J. Ethnopharmacol.*, vol. 210, pp. 296–310, Jan. 2018, doi: 10.1016/j.jep.2017.08.035.

[31]   S. Chen, V. Berthelier, J. B. Hamilton, B. O'Nuallai, and R. Wetzel, "Amyloid-like Features of Polyglutamine Aggregates and Their Assembly Kinetics," *Biochemistry*, vol. 41, no. 23, pp. 7391–7399, Jun. 2002, doi: 10.1021/bi011772q.

[32]   M. Oberstadt *et al.*, "TDP-43 self-interaction is modulated by redox-active compounds Auranofin, Chelerythrine and Riluzole," *Sci. Rep.*, vol. 8, no. 1, p. 2248, Feb. 2018, doi: 10.1038/s41598-018-20565-0.

[33]   T. Kaletta and M. O. Hengartner, "Finding function in novel targets: C. elegans as a model organism," *Nat. Rev. Drug Discov.*, vol. 5, no. 5, pp.387–399, May 2006, doi: 10.1038/nrd2031.

[34]   N. S. Tolwinski, "Introduction: Drosophila-A Model System for Developmental Biology," *J. Dev. Biol.*, vol. 5, no. 3, p. 9, Sep. 2017, doi: 10.3390/jdb5030009.

Human Neural Stem Cell Perilesional Transplant Potential Recovery in Penetrating Traumatic Brain Injury

Francesca Froio

Abstract— **Traumatic brain injury (TBI) is a leading cause of death and disability in the United States [1-5] Despite exceeding the death rate of cancer by 3.5 times, there is inadequate treatment that directly targets the TBI lesion, in particular the lesions due to penetrating traumatic brain injury (pTBI)[1]; pTBI is a niche area of TBI injuries that focuses on a foreign object entering and harming the brain [2]. Since the discovery of neural stem cells in the subventricular zone (SVZ) and dentate gyrus (DG), exploration of transplantation treatment has become a topic of interest [6-8] The aim of this study was to understand how stem cell treatments could be a optimized to address penetrating traumatic brain injury. It was initially thought that neural cells were non regenerative in central nervous system (CNS) injuries and that adult neurogenesis is limited in the SVZ and DG. However, neural stem cells are still present within the subventricular cortex after the injury. This demonstrates how transplanting endogenous cells could be a better treatment option in comparison to the current treatments that only mitigate secondary injuries and symptoms[3,9,10] Indeed, a growing number of experiments and animal trials have shown that human neural stem cells (hNSCs) transplanted perilesional to the cavity have the potential to aid pTBI recovery [2,5,11,12].**

I. INTRODUCTION

In the United States, around 1.7 million TBI cases are reported each year [13]. Specifically, pTBI contributes to the majority of firearm deaths [13]. That is about 20,000 headshots occur annually, and 70% of severe blast injuries result from pTBI [4,14]. Often pTBI leaves its patients with a lower life quality and long term disabilities such as Alzheimer's Disease, seizures, and neuroendocrine dysregulation. It also poses an economic burden costing $76.5 billion dollars for both indirect (loss in the workforce, emotional, psychosocial burdens, etc) and direct (emergency treatment, hospitalization, healthcare, etc.) expenses per year [13]. Despite this, there are presently no effective treatment methods for pTBI as current treatment of pTBI is primarily focused on managing secondary injury and symptoms [2,5,9,11,15].Therefore there is a need for more effective treatment methods that target the pTBI lesions directly. That is, therapies that are aimed at replacing the lost neurons within the resulting brain cavity.

Human neural stem cells hNSCs and their potential to promote proliferation (the increase in the number of such cells as a result of cell growth and division) and differentiation (a process in which young, unspecialized cells inhibit individual characteristics from their environment and adopt a specialized

function and form) within the resulting cavity, explicitly in the procephalic cyst, reveals a compelling future treatment option for the most important consequence of pTBI: neuronal loss [1,2,5-7,9,11,15,16,19-21]. Neurologists have demonstrated evidence of proliferation, differentiation, engraftment (growth of transplanted cells and successful interaction with new environment), reduced inflammation, and improvement of motor and cognitive deficits post hNSC transplant [2,5,9,11,15].

To explore how hNSCs could be an improved treatment option for the lost neurons in pTBI, primary and review journals demonstrating existing model hNSC transplantation were examined. Twenty plus journals from sources, such as PubMed, fell within the last decade of research. Key search words incorporated in the research process included perilesional, hNSC, pTBI, transplant, degeneration, and TBI in general. After organizing the data in accordance to the paper outline, original figures were curated utilizing Biorender. While research and primary experiments on the subject are limited, present data suggests that hNSCs may be a viable treatment option for pTBI.

II. PENETRATING TRAUMATIC BRAIN INJURY
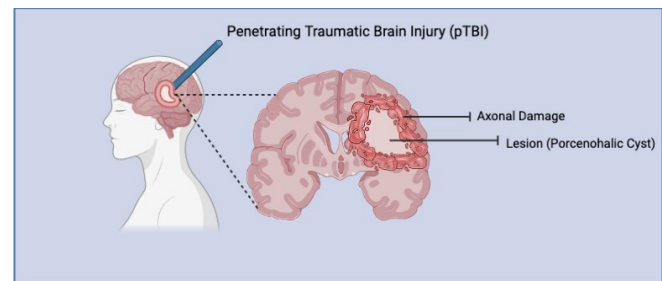
Figure 1: Representative of a probe or projectile penetrating the brain. Surrounding the lesion illustrates secondary damage at the cellular level, such as apoptosis, axonal damage, demyelination, and the formation of a porencephalic cyst (created using Biorender).

pTBI is defined as when an object breaches the skull, dura, and damages the parenchyma. Roughly, pTBI includes all traumatic brain injuries other than blunt head trauma (see Figure 1) [2]. Generally, when the projectile travels through the brain parenchyma, it causes a transient sonic wave which crushes the soft brain tissue and cultivates a permanent track of injury [3]. The severity of pTBI is heavily dependent on the velocity of the object at the point of penetration [3,6]. High velocity penetration consists of injuries produced by bullets or

shell fragments from direct trauma or shockwave injury surrounding brain tissue. On the other hand, low velocity penetration derives from sharp objects, such as a knife, causing direct trauma to brain tissue [2,16]. Besides velocity, the injury caliber is also determined by the intracranial path and energy/speed of object entry [3,14]. pTBI that contacts across the midline, passes through ventricles, or comes to rest in the posterior fossa, holds higher mortality rates in addition to projectiles maintaining higher velocity [3,14]. Consequently, the projectiles also determines the primary and secondary injury's austerity. Physics aside, other external factors determine the injury's severity. For instance, old age typically leads to a worse prognosis along with the pTBIs resulting from suicide attempts (due to closer proximity) [3].

III. PRIMARY AND SECONDARY INJURIES RELATED TO pTBI

Common primary injuries of pTBI include hemorrhage (blood loss), hematomas (blood clotted tissue), and parenchymal contusions (bruise of brain tissue) [1-3]. The risk of local wound infections, meningitis, ventriculitis, or cerebral abscess is also particularly high in pTBI patients due to contaminated foreign objects, skin, hair, and bone fragments along the projectile track [3-14]. But, one particularly evident resulting injury would be a porencephalic cyst (PC). When an object permeates the brain, a cavity typically results within the cerebral hemisphere. A PC is common in pTBI patients due to inflammation from the limited pool of hNSCs.[2] If CSF fills the cavity and affects the brain's communication with the ventricular system, this indicates the presence of a PC and is followed by a diagnosis of porencephaly, which is verified by a computed tomography (CT) scan [17]. Symptoms of a PC include visual field defects and brain behavior mimicking the presence of a stroke or brain tumor.

IV. CURRENT TREATMENT

A patient with pTBI is managed by a medical team that takes note of key information such as duration of loss of consciousness, seizures at any point in time, comorbidity (meaning the simultaneous presence of two diseases or conditions in a patient), and if any anticoagulants or antiplatelets (substances used to prevent and treat blood clots) were used [14].This early activation from the trauma team may aid in providing recognition of polytrauma (severely injured patients usually with two or more severe injuries in at least two areas of the body and an accurate severity assessment, considering the entry and exit points of the injury [3,14]. Primary analysis utilizing various neuroimaging techniques aid professionals in the evaluation and prognosis of pTBI. The Glasgow coma scale (GCS) is commonly taken of pTBI patients to scan for intracranial pressure [3,14]. A CT scan is also taken of a pTBI patient to evaluate the mass lesion or cerebral edema along with identifying the extent of any intracranial injury [3,14]. As this brief explanation of the process is undertaken in the current treatment indicates, there is a strong emphasis on resolving secondary injuries and symptoms rather than neuronal loss or axonal damage.

V. STEM CELLS

Stem cells (SC) are unspecialized, pluripotent cells, or cells that have the ability to give rise to any other type of cell within the body [7,8,18]. SCs differ from mature cells as mature cells are specialized and maintain a set function while SC's pluripotency makes them unique [14]. In order to be characterized as a stem cell, the cell must be able to self-renew (produce new stem cells) and differentiate (specialize into a specific cell type) [18]. Different classifications of stem cells affect these properties, as some stem cells might be multipotent and only give rise to cells in a specific family such as blood cells or totipotent and can form all cell types [8,11].

VI. EMBRYONIC STEM CELLS

Embryonic stem cells (ESC), while they are justifiable for TBI treatment as they maintain established protocols for maintenance in culture and are pluripotent (or able to give rise to any type of cell), are not researched thoroughly in current animal models of pTBI compared to hNSC [20]. Furthermore, generation of ESC is insufficient, unsure whether they would be rejected if used in transplants [20]. Therapies that use ESC lack concrete results and if derived directly from ESC undifferentiated culture prep can cause tumors and promote cancer development [20]. Applying hNSCs would avoid potential ethical issues associated with cell harvesting along with their multipotency with respect to differentiation into multiple neural phenotypes.

Author	Experimental Model	Significant Results
Hu et al.	Perilesional vs Intralesional transplant of hNSC in pTBI.	Perilesional mitigated secondary injury and reduced lesion size significantly compared to intralesional. Perilesional group also had signifigantly more motor cortex sparing and not persistance of a porcenphalic cyst, as evident in intralesional group.
Spurlock . M et al	Neuronal differentiation of hNSC post PBBI.	hNSC survived for at least 5 months. 150%of cells successfullt engrafted with 57% differentiation. Presence of presynaptic structural protein indicates integration with existing neural network.
Blaya M. et al	Genetically engineered MNS1 in hNSC transplantation post TBI.	Modified hNSC exhibited elevated survival rates. GFP/NeuN-positive cells doubled in engraftment in TBI MNS1 group compared to sham control. TBI MNS1 group had significantly greater hippo-campal neurogenesis and improved hippocampal dependent spatial memory capacity.
Elias P et al.	Implantation of Collagen scaffold seeded with hippocamapl models with pTBI.	Engraftment was low since scaffold didn't fill cavity. 1-2% cells remained in scaffold and BrdU survival was low. Positive astrocytes were detected around glial scarring.
Ruppert K et al.	CCI injured rats with transplanted MSC.	Significant increase of neurological deficits in early treatment. Delayed treatment demonstrated significant decrease in cognitive and neurological deficits as well as support for anti-inflammatory pathways.
Haus D et al.	hNSC transplant into immunodeficient rats with TBI.	hNSC groups preformed significantly greater in MWM test compared to sham. Demonstrated ameliorated impairment with improvement in learning locations. Improved host hippocampal neuron survival without harming lesion or host neuronal network while exhibiting migration throughout the host brain, extending in ventral hippocampal and ventral/medial cortical regions.

Table 1. Experimental models of stem cell transplants in TBI.

Differentiation and function improvement have been present in hNSCs experimental models already. At the cellular level, hNSC were also able to differentiate into neuronal and glial lineages, mitigate axonal damage, recapitulate host neural pathways, improve host neuron activity patterns and migrate beyond the lesion location (see Table 1) [5,10,12,20,21].Within the proliferation induced by hNSC transplant, the expression of nestin (cytoskeletal intermediate filament initially characterized in neural stem cells) was high post pTBI in the animal hNSCs transplant models [5,11,16,21]. Aside from differentiation, at the cellular level, hNSCs have restored expression of plasticity related Arc in host tissue, which has a key role in synaptic plasticity and memory consolidation. These results indicate that hNSCs transplantation in pTBI had rebuilt neuronal function [16].. However, it is important to keep in mind that it takes 1-3 months for hNSCs to fully differentiate into neurons, which explains why in most studies, after one week, the implanted neurons remained rounded and undifferentiated [11].

VII. MESENCHYMAL STEM CELLS (MSC)

While they are multipotent and have easier accessibility (because they can be isolated from various tissues), MSC treatment is not aimed to replace lost neurons, which is the main objective in any CNS stem cell transplant [13,19].An existing study incorporating MSC confirms these concerns as the study's inconsistency with cell quantity injected revealed no long-term engraftment and survival issues [13,20].

hNSC engraftment has been demonstrated extensively in experimental models. For instance, hNSCs were able to achieve 90% engraftment while interacting with existing neural microenvironments and reduce astrological scarring (scars evident in brain tissue) [2,5,10,15,16,20]. Although, in certain cases, transplanted hNSCs did not have any significant effect on reducing axon damage, hNSCs presents the best possible option for pTBI treatment due to their verified engraftment, lesion size reduction, and improvement of cognitive and motor deficits in rat pTBI and TBI models (see Table 1) [2,5,9–13,15,2].

VIII. NEURAL STEM CELLS (NSCs)

NSCs influence neuroblast migration toward the injury site, number of residential neurons and glial cells, astrogliosis, and locomotor recovery [11]. . That being said, hNSCs perform better compared to NSC as rat NSC gave rise to 27% new neurons while hNSCs gave rise to 57%.[5] hNSCs have also illustrated extended migration and differentiation outside the damaged tissue in cortical areas, the blood brain barrier (BBB), into vascular and endothelial cells, the medial ipsilateral cortex, the contralateral corpus callousness, and surrounding brain tissue [2,5,10,12,20,21]. Engraftment of hNSC has also been recognized long term, surpassing 5 months at least post transplantation and differentiation into mature neurons, astrocytes, and oligodendrocytes [5,19,20].

IX. LOCATION OF STEM CELLS TRANSPLANTATION

Location of hNSC transplantation influences the engraftment rate, migration, and impact on spatial and physical improvement [11,21]. Post pTBI, the resulting cavity lacks structural support and promotes apoptosis and neuronal death rather then engraftment [6]. Minutes to a few months after the pTBI was formed, pro-inflammatory cytokines that mobilize immune and glial cells to the injury environment, causing edema, inflammation, and demyelination (damage to the myelin sheath that surrounds neuronal fibers) [11]. The natural microenvironment at the brain with a raw pTBI cavity is not suitable for optimal success of transplanted stem cells, hence why transplanting the cells around the lesion would produce greater recovery and has been proven to do so in existing rat pTBI models [11,20,2].

An intralesional transplantation refers to the hNSC being transplanted directly into the resulting injury or cavity. A perilesional transplantation indicates that the subject was inserted around the cavity [2]. Comparing the two hNSC methods within a Sprague Dawley rat model of pTBI, the results of lesion size and motor cortex sparing of the perilesional group were significantly greater compared to the intralesional (see Figure 2) [2]. The study's foot fault test measured physical and cognitive deficits post transplantation. While there was significant lesion reduction and cortex sparing between the two groups, the test revealed insignificant data between the two in engraftment and behavioral difference. This leaves a gap in reasoning since significant cortex sparing should evidently produce significant behavioral differences.[2] Nonetheless, the perilesional transplantation lead to greater tissue/cortex preservation and should continue being tested and evaluated moving forward [2,5,12].
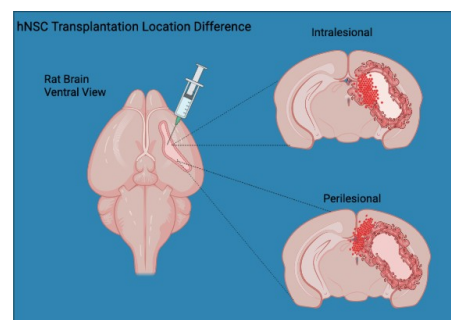
Figure 2. Perilesional transplantation compared to intralesional. Illustration of rat brain post experimental pTBI from the ventral view. Displaying the transplantation of hNSC, while demonstrating the difference between a perilesional and intralesional transplant in relation to the cavity (made using Biorender).

Another study confirmed the results previously presented as hNSCs reduced lesion size and increased neuronal differentiation through a perilesional transplantation.[5] Two main groups were observed: the sham or placebo group that had a mimicked pTBI but no cells transplanted (control group) and the transplant group that received the pTBI and cell transplantation. Even though latency was not significant between sham and transplant groups, this model conveyed the

perilesional transplant method reliable and viable, with 43% engraftment exhibited [5]. Previous animal models of transplanted hNSCs in other locations, such as contralaterally, or through a scaffold, left the cavity surrounded by glial scars, unfilled, and reduced migration (see Table 1) [10]. Present perilesional models have shown greater extent of differentiation and maturation [10].

X. FUTURE DIRECTIONS

Moving forward in hNSC model experimentation for pTBI treatments, a few factors should be considered. Primarily, a larger body of animal model experiments that evaluate engraftment and proliferation beyond two months should be conducted to work toward a potential human model. Along with engraftment, growth factors, and biogenic factors, promotional influences in differentiation should be considered further. Immunosuppression should also be acknowledged because of its critical nature in hNSC transplantation. As preclinical studies of TBI have generally established that hNSC transplantation was neuroprotective, the original lack of neuronal replacement is attributed to robust host immune system response rejection, which can be lessened through immunosuppression and promote greater engraftment in future studies.

XI. CONCLUSION

In summary, pTBI is a nationwide issue that could be ameliorated through hNSC in clinical practice. Current pTBI treatment, while good at managing secondary injury and symptoms, does not address the lost neurons nor the lesion head on. hNSC's potential to interact with host neural networks and restore neural connections effectively has been demonstrated through numerous animal models. By extending experiments past two months and observing long lasting activity of perilsionally transplanted hNSC, neurologists will be able to gauge how this treatment would function in human brains. Future studies should also focus on reducing the lesion along with behavioral and physical improvements, utilizing a wide variety of tests, for instance the Morris Water Maze test or the foot fault test to optimize progress. While further safety and mechanistic studies are warranted prior to the clinical trial phase, there is good evidence in support of a hNSC transplant as a treatment option for pTBI.

XII. REFRENCES

[1] Dekmak AS, Mantash S, Shaito A, et al. Stem cells and combination therapy for the treatment of traumatic brain injury. *Behavioural Brain Research*. 2018;340:49-62. doi:10.1016/J.BBR.2016.12.039

[2] Hu Z, Gajavelli S, Spurlock MS, et al. Human neural stem cell transplant location-dependent neuroprotection and motor deficit amelioration in rats with penetrating traumatic brain injury. In: *Journal of Trauma and Acute Care Surgery*. Vol 88. ; 2020. doi:10.1097/TA.0000000000002510

[3] Kazim SF, Shamim MS, Tahir MZ, Enam SA, Waheed S. Management of penetrating brain injury. *Journal of Emergencies, Trauma and Shock*. 2011;4(3). doi:10.4103/0974-2700.83871

[4] Shear DA, Lu XCM, Bombard MC, et al. Longitudinal characterization of motor and cognitive deficits in a model of penetrating ballistic-like brain injury. *Journal of Neurotrauma*. 2010;27(10). doi:10.1089/neu.2010.1399

[5] Spurlock MS, Ahmed AI, Rivera KN, et al. Amelioration of Penetrating Ballistic-Like Brain Injury Induced Cognitive Deficits after Neuronal

Differentiation of Transplanted Human Neural Stem Cells. *Journal of Neurotrauma*. 2017;34(11). doi:10.1089/neu.2016.4602

[6] Kassi AAY, Mahavadi AK, Clavijo A, et al. Enduring neuroprotective effect of subacute neural stem cell transplantation after penetrating TBI. *Frontiers in Neurology*. 2018;9. doi:10.3389/fneur.2018.01097

[7] Saha B, Jaber M, Gaillard A. Potentials of endogenous neural stem cells in cortical repair. *Frontiers in Cellular Neuroscience*. 2012;(MARCH). doi:10.3389/fncel.2012.00014

[8] Sun W, Kim H, Moon Y. Control of neuronal migration through rostral migration stream in mice. *Anatomy & Cell Biology*. 2010;43(4). doi:10.5115/acb.2010.43.4.269

[9] Blaya MO, Tsoulfas P, Bramlett HM, Dietrich WD. Neural progenitor cell transplantation promotes neuroprotection, enhances hippocampal neurogenesis, and improves cognitive outcomes after traumatic brain injury. *Experimental Neurology*. 2015;264. doi:10.1016/j.expneurol.2014.11.014

[10] Elias PZ, Spector M. Implantation of a collagen scaffold seeded with adult rat hippocampal progenitors in a rat model of penetrating brain injury. *Journal of Neuroscience Methods*. 2012;209(1). doi:10.1016/j.jneumeth.2012.06.003

[11] Dixon KJ, Theus MH, Nelersa CM, et al. Endogenous Neural Stem/Progenitor Cells Stabilize the Cortical Microenvironment after Traumatic Brain Injury. *Journal of Neurotrauma*. 2015;32(11). doi:10.1089/neu.2014.3390

[12] Wennersten A, Holmin S, al Nimer F, Meijer X, Wahlberg LU, Mathiesen T. Sustained survival of xenografted human neural stem/progenitor cells in experimental brain trauma despite discontinuation of immunosuppression. *Experimental Neurology*. 2006;199(2). doi:10.1016/j.expneurol.2005.12.035

[13] Ruppert KA, Prabhakara KS, Toledano-Furman NE, et al. Human adipose-derived mesenchymal stem cells for acute and sub-acute TBI. *PLoS ONE*. 2020;15(5). doi:10.1371/journal.pone.0233263

[14] Alao T. Penetrating head trauma . U.S. National Library of Medicine.

[15] Sullivan GM, Armstrong RC. Transplanted Adult Neural Stem Cells Express Sonic Hedgehog In Vivo and Suppress White Matter Neuroinflammation after Experimental Traumatic Brain Injury. *Stem Cells International*. 2017;2017. doi:10.1155/2017/9342534

[16] Acharya MM, Rosi S, Jopson T, Limoli CL. Human neural stem cell transplantation provides long-term restoration of neuronal plasticity in the irradiated hippocampus. *Cell Transplantation*. 2015;24(4). doi:10.3727/096368914X684600

[17] Yang DN, Townsend JC, Ilsen PF, Bright DC, Welton TH. Traumatic porencephalic cyst of the brain. *Journal of the American Optometric Association*. 1997;68(8). doi:10.1097/00006324-199512001-00062

[18] khanacademymedicine. Stem cells. Khan Academy .

[19] Zhou Y, Shao A, Xu W, Wu H, Deng Y. Advance of Stem Cell Treatment for Traumatic Brain Injury. *Frontiers in Cellular Neuroscience*. 2019;13. doi:10.3389/fncel.2019.00301

[20] Haus DL, López-Velázquez L, Gold EM, et al. Transplantation of human neural stem cells restores cognition in an immunodeficient rodent model of traumatic brain injury. *Experimental Neurology*. 2016;281. doi:10.1016/j.expneurol.2016.04.008

[21] Jahanbazi Jahan-Abad A, Sahab Negah S, Hosseini Ravandi H, et al. Human Neural Stem/Progenitor Cells Derived From Epileptic Human Brain in a Self-Assembling Peptide Nanoscaffold Improve Traumatic Brain Injury in Rats. *Molecular Neurobiology*. 2018;55(12). doi:10.1007/s12035-018-1050-8

Glioblastoma Multiforme: A Therapeutic Review

Mirika Jambudi

Abstract— **Glioblastoma multiforme (GBM) is one of the most common forms of malignant brain cancer. Despite advancements in technology and treatment over the past century, GBM remains largely incurable. Standard approaches for treatment include surgery and combinations of radiotherapy and chemotherapy, but factors such as the highly selective blood-brain barrier have made treating GBM and other brain diseases extremely difficult. However, immunotherapy or "personalized medicine" integrated with chemotherapy or radiotherapy may become the future for targeting GBM tumors and other brain diseases. This review evaluates the mechanisms and efficacy of standard drugs such as temozolomide and bevacizumab, as well as novel advancements in the field, such as nano-mediated drug delivery systems (NDDS) and the rise of immunology as a basis for treating GBM.**

I. INTRODUCTION

Accounting for more than 78% of brain cancers [1] and causing almost 15,000 deaths every year, glioblastoma multiforme (GBM) is one of the most common and aggressive malignant tumor forms in the central nervous system. GBM is characterized as a high-grade intra-axial tumor because it interferes with brain tissue [1]. Tumors are categorized as "low-grade" or "high-grade" depending on their invasiveness and growth rate [2], with low-grade cancers growing more slowly with less likelihood of metastasizing, or spreading to other sites of the body, than high-grade cancers [1]. GBM develops in glial cells, cells that protect neural tissue, causing a toxic buildup of glutamate, an excitatory neurotransmitter for cell signaling [3]. The excess glutamate kills surrounding neurons, creating brain space for the tumor to expand [4].

A variety of factors are taken into consideration when determining treatment, which may include varying combinations of surgery, radiation, and chemotherapy. Currently only two drugs, temozolomide and bevacizumab, are FDA-approved to treat GBM [5]. Unfortunately, these two chemotherapy drugs have had very limited impact on GBM patient survival rates [6]. Developing alternative and targeted therapies has posed a challenge as glioma tumor cells are protected by the blood-brain barrier (BBB), which is a highly selective semipermeable membrane that acts to protect the brain from pathogens and infections. Due to the barrier's highly selective permeability, many therapies are unable to cross this boundary [7]. This review will discuss traditional treatments and potentially new technology for the treatment of GBM.

II. TRADITIONAL TREATMENTS

One class of the oldest chemotherapy drugs used for GBM and other cancer forms is alkylating agents, which are able to permeate the BBB [8], making them an optimal choice for G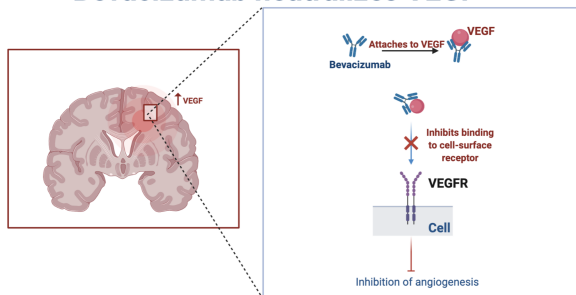BM treatment. Alkylating agents are used in cancer therapies due to their ability to prevent cells from replicating by inflicting damage to the cell's DNA [9]. Temozolomide (TMZ) is a typical alkylating agent used for GBM therapy, usually in conjunction with radiotherapy. The drug methylates DNA guanine bases [10], which results in alkylation of the DNA and DNA damage. Subsequently, this triggers apoptosis of malignant cells [11]. However, some tumor cells can become resistant to TMZ's effects, especially if the tumor cells have mutated and contain the gene MGMT that allows the cancerous cells to repair the DNA damage, preventing apoptosis and continuing the uncontrolled proliferation of the damaged cells [12]. Though TMZ-based chemotherapy demonstrates a comparable improvement in the treatment of patients who have high-grade gliomas, the median increase in survival for patients with GBM is only 2.5 months [13]. Recent studies also indicate that 60-75% of patients with GBM derive no benefit in regards to increased lifespan and quality of life from treatment with TMZ, demonstrating that the drug is only a modestly effective chemotherapy [13]. Additionally, 15-20% of patients who were treated with TMZ developed significant toxicity [14] and side effects such as amnesia and paralysis [15]. While TMZ is a widely-used drug, there is a significant need for chemotherapy or treatment with higher efficacy and safety.

Failure in treating GBM with TMZ chemotherapy led to the development of monoclonal antibodies and the introduction of targeted immunotherapies. Monoclonal antibodies have been used in therapy processes because of their high affinity for specific proteins involved in brain tumors and other cancers [16]. In 2004, the FDA approved bevacizumab (BEV), which inhibits angiogenesis, the development of new blood vessels, by neutralizing and blocking vascular endothelial growth factor (VEGF), a signaling protein that guides new vessel formation [18]. By targeting tumor growth mechanisms and inhibiting cell growth and division, BEV is able to block oncogenic signaling. Researchers have shown that glioma cells express and secrete VEGF, which has a positive correlation with increased tumor strength and aggressiveness. Since vascular proliferation is a hallmark of glioblastomas, [19], [20], BEV and its VEGF targeting mechanisms have been introduced for GBM.

With regards to GBM, bevacizumab slows tumor growth, but it does not cure the actual tumor itself or prolong overall patient survival time [16]. Additionally, rebound phenomena such as tumor recurrence and regrowth are often observed after discontinuation of BEV therapy [21]. Adverse side effects, such as hypertension and proteinuria are also associated with BEV usage [22]. While BEV has been shown to improve the quality of life for patients and has slight efficacy in recurrent GBM [23], [24], it is still only modestly

effective in treating GBM overall. With the need for more effective treatment, the basic mechanism of bevacizumab as a monoclonal antibody has led to the development of new

Figure 1: Bevacizumab is a monoclonal antibody that inhibits angiogenesis, the process by which new blood vessels form. It does so by blocking vascular endothelial growth factor (VEGF), a signaling protein that guides new vessel formation and is expressed in glioma cells. It has led to the development of new immunotherapies and advanced alternative treatment options, such as a vaccine targeting VEGF receptors in neurofibromatosis type 2 [62].

immunotherapies and advanced technology systems in treating GBM.

III. NOVEL TREATMENT OPTIONS

Though traditional drugs have had some limited success in treating GBM, nanotherapeutic drug delivery systems (NDDSs) and nanocarriers, transport vehicles for drugs, are rising in popularity as new alternative targeted cancer treatments. Compared to traditional drugs, NDDSs have been shown to have increased advantages when it comes to treating cancers, such as improved stability, enhanced permeability, and highly accurate targeting [25], [26]. Additionally, they have been shown to overcome cancer-related drug resistance by targeting resistance mechanisms including defective programmed cell death and overexpression of transporters [27]. Using NDDSs for treatment of brain cancer has become a promising alternative, as it is more effective at transporting chemotherapeutics across the BBB than traditional therapies and has minimal side effects on healthy, surrounding tissue [28], [29]. Dp44mt (Di-2-pyridyl ketone-4, 4-dimethyl-3-thiosemicarbazone) is a novel glioma-targeted nano-therapeutic that has been found to specifically target its toxicity towards glioma cells with no impact on the surrounding healthy tissue [28], [29]. When tested in mice, the Dp44mt nanoparticles reduced tumor growth by 62%. Other chemotherapies, such as TMZ and doxorubicin only reduced tumor growth by 16% [27], [30]–[32]. This may lead to better prognosis, and Dp44mt may serve as a more effective treatment for GBM in humans.

Attached to the nanocarrier, Dp44mt has a glioma-targeted ligand to Interleukin-13 (IL-13), which is found on gliomas [28]. In experimental studies, researchers found that Dp44mt's conjugation with IL-13 receptors on the tumor enhanced glioma cell uptake of the nanocarrier and allowed for more successful permeation of the BBB [28]. Dp44mt is

an iron chelator, which extracts excess iron from cells. Though iron is not the underlying cause of many diseases, it does play a role in the rate of disease progression through facilitation of cellular growth and proliferation [33]. For cancer cells, the chelator removes the iron they need for maintaining basic cellular functions, thus starving them [34]. Dp44mt, with the use of a nanocarrier, is the first instance of testing a nano-therapeutic system on brain tumors; it has yielded successful results, as this chelator has been able to overcome multidrug resistance, a common trait of high-grade tumors that renders them immune to chemotherapies [35].

While the drug is still undergoing numerous trials before reaching FDA consideration for approval, certain components of the drug, such as the nanoparticles used to create the carrier, have already been approved [29], [36].

Though a novel form of targeted therapy, nanocarriers and nanotherapeutic drug delivery systems hold promise for the future of cancer therapies. However, as this is still a technology undergoing preliminary testing, the drug's success in animal models may not translate completely to patients, and side effects are still unknown in humans. With the uncertainty surrounding this new technology combined with the low efficacy and adverse side effects of traditional treatments, research has found focus on personalized immunotherapy.

IV. RISE OF PERSONALIZED IMMUNOTHERAPIES

Vaccines are among the most standard forms of immunotherapy for bacteria and viruses. Now, vaccines are on the rise to treat diseases such as Alzheimer's and cancer [37], [38]. Some vaccines that prevent certain viral infections such as human papillomavirus (HPV) and hepatitis B have been modified to serve as cancer vaccines [37]. Due to this repurposing, vaccine therapy for GBM has risen in popularity with the study and development of vaccines in three primary categories: peptide, heat-shock, and cell-based [38]–[41]. Currently, a recent vaccine study for human epidermal growth factor receptor 2 (HER2)-positive breast cancer moved forward after successful results in preventing cancer reformation [42]. In addition to being expressed in breast cancer, upregulated expression of HER2 has been identified in GBM, and could potentially be an immunotherapy target [43]. With the preliminary success of the HER2 vaccine for breast cancer, it could potentially be used as an immunotherapy for GBM as well.

Several current Phase I and Phase II trials for GBM studying immunotherapies have shown tumor reduction and lifespan expansion, as 20% of patients in the study survived from four to five years, which is unusual considering the high fatality of GBM [44]. Compared to other forms of treatment, vaccine immunotherapies are compelling because they have minimal toxicity and can induce a highly patient-personalized anti-tumor response that may be key to eradicating GBM [40].

Additionally, as each vaccine is highly unique to each patient's immune system, it aligns with the upcoming

concept of "personalized medicine" [45], [47]. Personalized medicine is more effective than standard medication as treatment is tailored to the genes of each specific person [45], which has been shown to have high efficacy in cancers such as breast cancer [46], [48]. It may make GBM, one of the most malignant human tumors, manageable for patients while reducing side effects and increasing quality of life [48]. However, vaccine therapy does face some challenges, as surgical removal or biopsy of the tumor may be necessary in order to identify pathology and prepare the vaccine accordingly [49]. Furthermore, because each vaccine is individualized to each patient, this treatment method may not be affordable for all patients. However, as more advances in technology development and existing trials continue, manufacturers may find a cheaper way to create these vaccines. Though it may be an expensive treatment as of now, the precision of personalized medicine can improve the overall quality of life after therapy compared to other treatments, and the results outweigh the cost.

mRNA vaccines have also shown promise in regard to cancer immunotherapy. After vaccination, vehicle-loaded mRNA vaccines express tumor antigens in antigen-presenting cells (APCs), causing APC activation and stimulation of the innate and adaptive immune system [50]. mRNA cancer vaccines hold high promise over other vaccine forms due to their specific toxicity to tumor cells, increased safety, and cost-effectiveness [50], [51]. However, mRNA vaccines have had limitations such as instability in their ability to break down and inefficient delivery in vivo to tumor cells [52]. Nucleotide modifications and other alterations have been investigated to overcome these challenges, and numerous studies are underway [53]. There also is potential in repurposing treatments, such as the COVID-19 vaccine, to treat GBM. Combinations of mRNA vaccines with other immunotherapies may also increase the anti-tumor immune response. With the recent FDA approval of mRNA vaccines for COVID-19 and promising results of other mRNA cancer vaccines against aggressive solid tumors [51], mRNA vaccines may be a potential immunotherapy treatment for GBM.

Though mainstream therapies have had limited success and other forms of immunotherapies are still undergoing trials, the development of chimeric antigen reporter (CAR) T-cell therapy has also shown promise in treating GBM [54]. The treatment relies on using the patients' collected and genetically engineered cells targeting specific tumor-associated antigens [55]. These cells are harvested from the patient, modified to target particular proteins that the tumor expresses, then injected into the patient to destroy the tumor cells [56]. Once the CAR construct binds to the intended target antigen, the T cells are activated, prompting a cytokine release [57]. CAR T has been approved for use in other cancers, such as acute lymphoblastic leukemia and non-Hodgkin's lymphoma [59]. Complete remission rates for patients with leukemia undergoing CAR T therapy have been as high as 68%-93%, indicating the treatment's high efficacy and potential [59], [60]. The approach used in these other

cancers is now being applied to treating GBM [59]. There has been evidence that CAR T cells injected directly into the brain tumor tissue or spinal fluid may cause positive responses in patients [60], though a clinical trial is still underway for results to be validated.

The efficacy of CAR T therapy is still yet to be determined in GBM, as only preliminary studies of CAR T in GBM have been conducted. Therefore, it is essential to continue studying CAR T in the context of GBM since prior cancer studies have shown CAR T's effectiveness as a treatment option. Its application to GBM is still limited due to the lack of identified tumor-specific antigens expressed in the disease [61]. However, further advances in CAR T, such as multitargeting CAR T therapy, may be effective. Targets such as HER2, IL-13, and EGFRvIII have been identified as antigens involved in GBM, but there are numerous other antigens that have yet to be explored [55].

V. CONCLUSION

GBM has been one of the solid tumor cancers that are the most difficult to treat, despite advances in recent technology and medicine. Current standards of care, such as TMZ and radiotherapy, have had limited success in treating patients, often resulting in a myriad of side effects that can be fatal, as well as a significant decrease in the quality of life for patients. As GBM is notoriously difficult to treat due to its high aggressiveness, there is a significant need for treatments with higher efficacy and safety.

Immunotherapy has emerged as a promising choice for treatment, alongside the concept of "personalized medicine." With numerous treatments under development or undergoing studies and trials, immunotherapies such as vaccines for GBM and CAR T therapy have shown positive results in efficacy, as well as reduced side effects.

This review discussed standard forms of treatment and introduced a new perspective regarding the rise of novel immunotherapies for use in GBM, including vaccines and CAR T. With their revolutionary success in treating other diseases, these therapies have significant potential for GBM. While this review does not have an exhaustive list of therapies, it provides insight into novel therapeutics, building off of the standard treatments currently available.

Based on the direction that these immunotherapies are taking, there is a significant likelihood that future clinical trials will place a greater emphasis on efficacy, safety, immune system mechanisms, and drug resistance prevention. With this, the future of GBM may be combinations of CAR T therapy, vaccines, and other modes of standard treatment, such as chemotherapy, radiation, surgery, etc., making the modern concept of "personalized medicine" a reality.

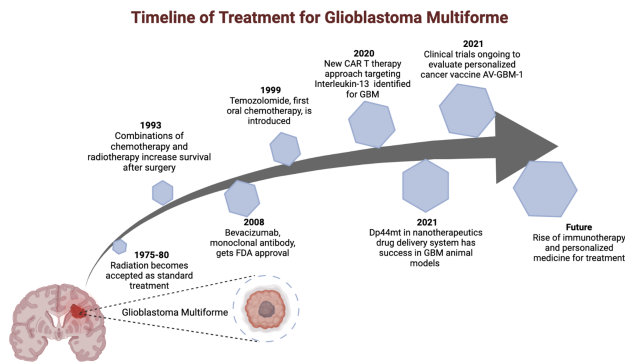**Timeline of Treatment for Glioblastoma Multiforme**

Figure 2: Radiation therapy and chemotherapy have been accepted as traditional forms of treatment for GBM but are still not sufficient. The rise of immunotherapy and "personalized medicine" have led to the development of potential new technology for the treatment of GBM, many of which are undergoing clinical trials and testing.

VI. Acknowledgments

I would like to thank my mentor Nicole Katchur for assisting me with my project's development, as well as my family for their continuous support.

VII. References

[1] Wirsching, H. G., Galanis, E., & Weller, M. (2016). Glioblastoma. *Handbook of Clinical Neurology*, *134*. https://doi.org/10.1016/B978-0-12-802997-8.00023-2

[2] Kanderi, T., & Gupta, V. (2021). Glioblastoma Multiforme. In *StatPearls [Internet]*. StatPearls Publishing.

[3] Siva Kumar Natarajan, S. V. (2019). Glutamine Metabolism in Brain Tumors. *Cancers*, *11*(11). https://doi.org/10.3390/cancers11111628

[4] *The Love-Hate Relationship with Glial Cells - Science in the News*. (2008, June 16). https://sitn.hms.harvard.edu/flash/2008/issue43/

[5] Jacob P. Fisher, D. C. A. (2021). Current FDA-Approved Therapies for High-Grade Malignant Gliomas. *Biomedicines*, *9*(3). https://doi.org/10.3390/biomedicines9030324

[6] Kang, Y. J., Holley, C. K., Abidian, M. R., Madhankumar, A. B., Connor, J., & Majd, S. (2020). Tumor targeted delivery of an anti-cancer therapeutic: An in vitro and in vivo evaluation. Advanced Healthcare Materials, 10(2), 2001261. https://doi.org/10.1002/adhm.202001261

[7] Richard Daneman, A. P. (2015). The Blood–Brain Barrier. *Cold Spring Harbor Perspectives in Biology*, *7*(1). https://doi.org/10.1101/cshperspect.a020412

[8] Pardridge, W. M. (2012). Drug transport across the blood–brain barrier. *Journal of Cerebral Blood Flow and Metabolism: Official Journal of the International Society of Cerebral Blood Flow and Metabolism*, *32*(11), 1959.

[9] Colvin, M. (2003). Alkylating Agents. In *Holland-Frei Cancer Medicine. 6th edition*. BC Decker.

[10] Zhang, J., Stevens, M. F., & Bradshaw, T. D. (2012). Temozolomide: mechanisms of action, repair and resistance. *Current Molecular Pharmacology*, *5*(1). https://doi.org/10.2174/1874467211205010102

[11] Barciszewska, A. M., Gurda, D., Głodowicz, P., Nowak, S., & Naskręt-Barciszewska, M. Z. (2015). A New Epigenetic Mechanism of Temozolomide Action in Glioma Cells. *PloS One*, *10*(8). https://doi.org/10.1371/journal.pone.0136669

[12] Wesolowski, J. R., Rajdev, P., & Mukherji, S. K. (2010). Temozolomide (Temodar). *AJNR. American Journal of Neuroradiology*, *31*(8), 1383–1384.

[13] Chamberlain, M. C. (2010). Temozolomide: therapeutic limitations in the treatment of adult high-grade gliomas. *Expert Review of Neurotherapeutics*, *10*(10). https://doi.org/10.1586/ern.10.32

[14] Perry, J. R., Bélanger, K., Mason, W. P., Fulton, D., Kavan, P., Easaw, J., Shields, C., Kirby, S., Macdonald, D. R., Eisenstat, D. D., Thiessen, B., Forsyth, P., & Pouliot, J. F. (2010). Phase II trial of continuous dose-intense temozolomide in recurrent malignant glioma: RESCUE study. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, *28*(12). https://doi.org/10.1200/JCO.2009.26.5520

[15] Bae, S. H., Park, M.-J., Lee, M. M., Kim, T. M., Lee, S.-H., Cho, S. Y., Kim, Y.-H., Kim, Y. J., Park, C.-K., & Kim, C.-Y. (2014). Toxicity Profile of Temozolomide in the Treatment of 300 Malignant Glioma Patients in Korea. *Journal of Korean Medical Science*, *29*(7), 980.

[16] Ameratunga, M., Pavlakis, N., Wheeler, H., Grant, R., Simes, J., Khasraw, M., Gynaecological, C., Neuro-oncology, & Orphan Cancer Group. (2018). Anti-angiogenic therapy for high-grade glioma. *Cochrane Database of Systematic Reviews*, *2018*(11). https://doi.org/10.1002/14651858.CD008218.pub4

[17] Iwamoto, F. M., & Fine, H. A. (2010). Bevacizumab for Malignant Gliomas. *Archives of Neurology*, *67*(3), 285–288.

[18] Mukherji, S. K. (2010). Bevacizumab (Avastin). *AJNR. American Journal of Neuroradiology*, *31*(2), 235–236.

[19] Li, Y., Ali, S., Clarke, J., & Cha, S. (2017). Bevacizumab in Recurrent Glioma: Patterns of Treatment Failure and Implications. *Brain Tumor Research and Treatment*, *5*(1), 1.

[20] Ascha, M. S., Wang, J. F., Kumthekar, P., Sloan, A. E., Kruchko, C., & Barnholtz-Sloan, J. S. (2019). Bevacizumab for the treatment of non-small cell lung cancer patients with synchronous brain metastases. *Scientific Reports*, *9*(1), 1–9.

[21] Narita, Y. (2013). Drug Review: Safety and Efficacy of Bevacizumab for Glioblastoma and Other Brain

Tumors. *Japanese Journal of Clinical Oncology*, *43*(6), 587–595.

[22] Gil-Gil, M. J., Mesia, C., Rey, M., & Bruna, J. (2013). Bevacizumab for the Treatment of Glioblastoma. *Clinical Medicine Insights. Oncology*, *7*, 123.

[23] De Fazio, S., Russo, E., Ammendola, M., Di Paola E, D., & De Sarro, G. (2012). Efficacy and safety of bevacizumab in glioblastomas. *Current Medicinal Chemistry*, *19*(7). https://doi.org/10.2174/092986712799320646

[24] Yu, Z., Zhao, G., Zhang, Z., Li, Y., Chen, Y., Wang, N., Zhao, Z., & Xie, G. (2016). Efficacy and safety of bevacizumab for the treatment of glioblastoma. *Experimental and Therapeutic Medicine*, *11*(2), 371.

[25] Yao, Y., Zhou, Y., Liu, L., Xu, Y., Chen, Q., Wang, Y., Wu, S., Deng, Y., Zhang, J., & Shao, A. (2020). Nanoparticle-Based Drug Delivery in Cancer Therapy and Its Role in Overcoming Drug Resistance. *Frontiers in Molecular Biosciences*, *0*. https://doi.org/10.3389/fmolb.2020.00193

[26] Jain, K. K. (2007). Use of nanoparticles for drug delivery in glioblastoma multiforme. *Expert Review of Neurotherapeutics*, *7*(4). https://doi.org/10.1586/14737175.7.4.363

[27] Gallego, L., & Ceña, V. (2020). Nanoparticle-mediated therapeutic compounds delivery to glioblastoma. *Expert Opinion on Drug Delivery*, *17*(11). https://doi.org/10.1080/17425247.2020.1810015

[28] Holley, C. K., & Majd, S. (2020). Examining the Anti-Tumor Activity of Dp44mT-Loaded Nanoparticles In Vitro. *Conference Proceedings: ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, *2020*. https://doi.org/10.1109/EMBC44109.2020.9176197

[29] Zhou, J., Jiang, Y., Zhao, J., Zhang, H., Fu, J., Luo, P., Ma, Y., Zou, D., Gao, H., Hu, J., Zhang, Y., & Jing, Z. (2020). Dp44mT, an iron chelator, suppresses growth and induces apoptosis via RORA-mediated NDRG2-IL6/JAK2/STAT3 signaling in glioma. *Cellular Oncology: The Official Journal of the International Society for Cellular Oncology*, *43*(3), 461–475.

[30] Alimohammadi, E., Bagheri, S. R., Taheri, S., Dayani, M., & Abdi, A. (2020). The impact of extended adjuvant temozolomide in newly diagnosed glioblastoma multiforme: a meta-analysis and systematic review. *Oncology Reviews*, *14*(1). https://doi.org/10.4081/oncol.2020.461

[31] Liao, W.-H., Hsiao, M.-Y., Kung, Y., Huang, A. P.-H., & Chen, W.-S. (2021). Investigation of the Therapeutic Effect of Doxorubicin Combined With Focused Shockwave on Glioblastoma. *Frontiers in Oncology*, *0*. https://doi.org/10.3389/fonc.2021.711088

[32] Da Ros, M., Iorio, A. L., De Gregorio, V., Fantappiè, O., Laffi, G., de Martino, M., Pisano, C., Genitori, L., & Sardi, I. (2018). Aldoxorubicin and Temozolomide combination in a xenograft mice model of human glioblastoma. *Oncotarget*, *9*(79), 34935.

[33] Hatcher, H. C., Singh, R. N., Torti, F. M., & Torti, S. V. (2009). Synthetic and natural iron chelators: therapeutic potential and clinical use. *Future Medicinal Chemistry*, *1*(9). https://doi.org/10.4155/fmc.09.121

[34] Cao, L. L., Liu, H., Yue, Z., Liu, L., Pei, L., Gu, J., Wang, H., & Jia, M. (2018). Iron chelation inhibits cancer cell growth and modulates global histone methylation status in colorectal cancer. *Biometals: An International Journal on the Role of Metal Ions in Biology, Biochemistry, and Medicine*, *31*(5). https://doi.org/10.1007/s10534-018-0123-5

[35] Holley, C. K., Alkhalifah, S., & Majd, S. (2018). Fabrication and Optimization of Dp44mT-Loaded Nanoparticles. *Conference Proceedings: ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, *2018*. https://doi.org/10.1109/EMBC.2018.8513598

[36] Mirab, F., Kang, Y. J., & Majd, S. (2019). Preparation and characterization of size-controlled glioma spheroids using agarose hydrogel microwells. *PloS One*, *14*(1). https://doi.org/10.1371/journal.pone.0211078

[37] Mamai, O., Dodagatta-Marri, E., & Akhurst, R. J. (2018). From prevention to cure, repurposing antiviral vaccines for cancer immunotherapy. *Biotarget*, *2*. https://doi.org/10.21037/biotarget.2018.12.03

[38] Thomas, A. A., Fisher, J. L., Ernstoff, M. S., & Fadul, C. E. (2013). Vaccine-based immunotherapy for glioblastoma. *CNS Oncology*, *2*(4). https://doi.org/10.2217/cns.13.29

[39] Butterfield, L. H. (2016). Lessons learned from cancer vaccine trials and target antigen choice. *Cancer Immunology, Immunotherapy: CII*, *65*(7). https://doi.org/10.1007/s00262-016-1801-1

[40] Xu, L. W., Chow, K. K., Lim, M., & Li, G. (2014). Current vaccine trials in glioblastoma: a review. *Journal of Immunology Research*, *2014*. https://doi.org/10.1155/2014/796856

[41] Swartz, A. M., Shen, S. H., Salgado, M. A., Congdon, K. L., & Sanchez-Perez, L. (2018). Promising vaccines for treating glioblastoma. *Expert Opinion on Biological Therapy*, *18*(11). https://doi.org/10.1080/14712598.2018.1531846

[42] Knutson, K. L., Block, M. S., Norton, N., Erskine, C. L., Hobday, T. J., Dietz, A. B., & Degnim, A. C. (2020). Rapid Generation of Sustainable HER2-specific T-cell Immunity in Patients with HER2 Breast Cancer using a Degenerate HLA Class II Epitope Vaccine. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, *26*(5), 1045–1053.

[43] Haynik, D. M., Roma, A. A., & Prayson, R. A. (2007). HER-2/neu Expression in Glioblastoma Multiforme. In *Applied Immunohistochemistry & Molecular Morphology* (Vol. 15, Issue 1, pp. 56–58). https://doi.org/10.1097/01.pai.0000213133.09160.da

[44] Kong, Z., Wang, Y., & Ma, W. (2018). Vaccination in the immunotherapy of glioblastoma. *Human Vaccines & Immunotherapeutics*, *14*(2), 255.

[45] Li, J., Di, C., Mattox, A. K., Wu, L., & Adamson, D. C. (2010). The future role of personalized medicine in the treatment of glioblastoma multiforme. *Pharmacogenomics and Personalized Medicine*, *3*. https://doi.org/10.2147/PGPM.S6852

[46] Sabatier, R., Gonçalves, A., & Bertucci, F. (2014). Personalized medicine: present and future of breast cancer management. *Critical Reviews in Oncology/hematology*, *91*(3). https://doi.org/10.1016/j.critrevonc.2014.03.002

[47] Jeibouei, S., Akbari, M. E., Kalbasi, A., Aref, A. R., Ajoudanian, M., Rezvani, A., & Zali, H. (2019). Personalized medicine in breast cancer: pharmacogenomics approaches. *Pharmacogenomics and Personalized Medicine*, *12*, 59.

[48] Taghizadeh, H., Müllauer, L., Furtner, J., Hainfellner, J. A., Marosi, C., Preusser, M., & Prager, G. W. (2019). Applied Precision Cancer Medicine in Neuro-Oncology. *Scientific Reports*, *9*(1), 1–8.

[49] Cuoco, J. A., Benko, M. J., Busch, C. M., Rogers, C. M., Prickett, J. T., & Marvin, E. A. (2018). Vaccine-Based Immunotherapeutics for the Treatment of Glioblastoma: Advances, Challenges, and Future Perspectives. *World Neurosurgery*, *120*. https://doi.org/10.1016/j.wneu.2018.08.202

[50] Miao, L., Zhang, Y., & Huang, L. (2021). mRNA vaccine for cancer immunotherapy. *Molecular Cancer*, *20*(1). https://doi.org/10.1186/s12943-021-01335-5

[51] Pardi, N., Hogan, M. J., Porter, F. W., & Weissman, D. (2018). mRNA vaccines — a new era in vaccinology. *Nature Reviews. Drug Discovery*, *17*(4), 261.

[52] Tang, X., Zhang, S., Fu, R., Zhang, L., Huang, K., Peng, H., Dai, L., & Chen, Q. (2019). Therapeutic Prospects of mRNA-Based Gene Therapy for Glioblastoma. *Frontiers in Oncology*, *9*. https://doi.org/10.3389/fonc.2019.01208

[53] Wang, Y., Zhang, Z., Luo, J., Han, X., Wei, Y., & Wei, X. (2021). mRNA vaccine: a potential therapeutic strategy. *Molecular Cancer*, *20*(1), 1–23.

[54] Bagley, S. J., Desai, A. S., Linette, G. P., June, C. H., & O'Rourke, D. M. (2018). CAR T-cell therapy for glioblastoma: recent clinical advances and future challenges. *Neuro-Oncology*, *20*(11). https://doi.org/10.1093/neuonc/noy032

[55] Maggs, L., Cattaneo, G., Dal, A. E., Moghaddam, A. S., & Ferrone, S. (2021). CAR T Cell-Based Immunotherapy for the Treatment of Glioblastoma. *Frontiers in Neuroscience*, *0*. https://doi.org/10.3389/fnins.2021.662064

[56] Waldman, A. D., Fritz, J. M., & Lenardo, M. J. (n.d.). A guide to cancer immunotherapy: from T cell basic science to clinical practice. *Nature Reviews. Immunology*, 1.

[57] Miliotou, A. N., & Papadopoulou, L. C. (2018). CAR T-cell Therapy: A New Era in Cancer Immunotherapy. *Current Pharmaceutical Biotechnology*, *19*(1). https://doi.org/10.2174/1389201019666180418095526

[58] Bupha-Intr, O., Haeusler, G., Chee, L., Thursky, K., Slavin, M., & Teh, B. (2021). CAR-T cell therapy and infection: a review. *Expert Review of Anti-Infective Therapy*, *19*(6). https://doi.org/10.1080/14787210.2021.1855143

[59] Land, C. A., Musich, P. R., Haydar, D., Krenciute, G., & Xie, Q. (2020). Chimeric antigen receptor T-cell therapy in glioblastoma: charging the T cells to fight. *Journal of Translational Medicine*, *18*(1), 1–13.

[60] Akhavan, D., Alizadeh, D., Wang, D., Weist, M. R., Shepphird, J. K., & Brown, C. E. (2019). CAR T cells for brain tumors: Lessons learned and road ahead. *Immunological Reviews*, *290*(1). https://doi.org/10.1111/imr.12773

[61] Karschnia, P., Teske, N., Thon, N., Subklewe, M., Tonn, J.-C., Dietrich, J., & von Baumgarten, L. (2021). Chimeric Antigen Receptor T Cells for Glioblastoma. Neurology, 97(5), 218–230.

[62] Tamura, R., Fujioka, M., Morimoto, Y. et al. A VEGF receptor vaccine demonstrates preliminary efficacy in neurofibromatosis type 2. Nat Commun 10, 5758 (2019). https://doi.org/10.1038/s41467-019-13640-1