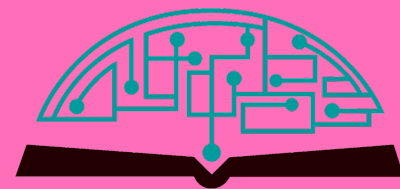


IJHSR

International
Journal of
High School
Research



September 2026 | Volume 8 | Issue 17

ijhsr.terrajournals.org

ISSN (Print) 2642-1046

ISSN (Online) 2642-1054



Let's build a better future together

International Project Fair focused on Sustainability and Environment for Grades 8-12



- Since 2011
- Hosted around 1400 participants in 2025 from 35+ states 70+ countries
- **Disciplines:** STEM, Coding, Robotics, AI, Speech, Entrepreneurship, Arts, Short Film, Music
- Applications start on December 1
- Application Deadline is March 1
- Finalists are announced by March 25
- Event is usually scheduled 2nd week of June
- Monday – Friday, includes a trip to Niagara Falls
- Hosted by large universities at Upstate New York
- Application Fee is \$75/ project
- Participation Fee is \$695/ person, w/ room and board
- Open buffet breakfast, lunch, and dinner
- Trip to Niagara Falls and boat tour is included
- **Instagram and Facebook @Geniusolympiad**
- For more information: **GENIUSOlympiad.org**
- Email: **info@geniusolympiad.org**

GENIUS Olympiad is organized by Terra Science and Education, a N.Y. based 501.c.3 non-profit organization dedicated for project-based learning

Table of Contents

September 2026 | Volume 8 | Issue 17

1	Using a Deep Neural Network to Downscale a General Circulation Model at the National Scale in Australia <i>Alex H. Zhuang</i>
8	Machine Learning in Predicting Outcomes of Brain Tumor Resection Surgeries: A Scholarly Review <i>Arman Raoufian</i>
14	Quantitative Analysis of Per- and Polyfluoroalkyl Substances (PFAS) in Selective South Bay Water Bodies <i>Aryanka V. Rekbi</i>
19	Evaluation of Nitrate Concentration Estimation using Machine Learning Models in Lake Erie Watersheds, USA <i>Saeyool Chloe Park</i>
24	Adjusting Valuation Metrics for the Intangible Economy: The Impact of R&D Capitalization on Microsoft and Amazon 2012–22 <i>Claudia Fanelli</i>
31	An Intuitive and General Learning Approach for Function Transformations <i>Elizabeth Kong</i>
39	Etomidate-Induced Cortical Inhibition: Prefrontal–Amygdala and Its Implications for Emotion and Behavior <i>Joshua Goh</i>
46	Emerging Downstream Plastic Waste Disposal Solutions in the United States: A Review of Plastic-Modified Pavement <i>Jasmine S. Liu</i>
64	Experiences of Compassion Fatigue in Doctors: An Exploration of Triggers, Coping Strategies, and Culture <i>Lea C. Y. Yuen</i>
72	Media Sentiment and Economic Signals of Emerging Blockchain Adoption in Vietnam <i>Gia Minh (Michael) Pham</i>
85	Enhanced Combination Therapy of Aminoglycosides Against Resistant <i>Staphylococcus aureus</i> <i>Olivia Liu</i>
92	A Systematic Review of Takotsubo Cardiomyopathy in Oncologic Cohorts <i>Pranav Maddineni</i>
106	The Role of Amyloid-β and Tau Proteins in Alzheimer's Disease <i>Pranavi Nandyala, Jobin Varkey</i>
114	Assessing the Impact of Microplastics on DNA Damage in <i>C. elegans</i> <i>Renee Ma</i>
120	How To Solve a Rubik's Cube: Using Search Algorithms <i>Richard Y. Tao</i>
124	Efficiency of Lipid Nanoparticles in Glioblastoma Therapies <i>Saimanasaa Viswanathan</i>
131	Low Self-Esteem and the School Environment: A Comparative Study of the Risk Factors Among Adolescents <i>Samarth Kasaraneni, Trayi Yada, Samuel Xiao</i>
136	Beyond the Pale: How Cognitive Biases Influence Our Preference for White Tree-based Paper <i>Seoba Han</i>
142	CRISPR-Cas9 for Treating Genetic Blood Disorders: Promises and Limits <i>Shriyans Katukuri</i>
150	Suitability of SAPARD Methods in the Post-War Recovery of Ukraine's Agricultural Economy <i>Sofiia Yanchuk</i>
160	Mathematical and Machine Learning-Based Modeling of Plant Recovery Dynamics Under Environmental Stress <i>Yenula Wickramasinghe</i>
165	Orexin-Induced Upregulation of <i>OX1R</i> And <i>BDNF</i> Suggests an Adaptive Neuroprotective Mechanism Against Oxidative Stress <i>Youn Joo Lee</i>
170	A Pilot Investigation of Antibiotic Susceptibility Profile of <i>Escherichia coli</i> in Wastewater from a Boarding High School in Türkiye <i>Zeynep M. Çelik</i>

Using a Deep Neural Network to Downscale a General Circulation Model at the National Scale in Australia

Alex H. Zhuang

The King's School, 87-129 Pennant Hills Road, North Parramatta, NSW, 2159, Australia; zhuanghuijie2008@gmail.com

Mentor: Joe Xiao

ABSTRACT: General Circulation Models (GCMs) are powerful tools for simulating future climate change, but have suffered from low-resolution outputs that are unusable for local and regional planning and budgetary bodies. This study aims to downscale GCM-simulated annual rainfall data at the national scale in Australia, representing the largest-scale GCM downscaling study in Australian Literature. This study collected over 69000 data points containing hydrological variables and mapped them within GIS software. These data points were arranged in a grid with a spacing of 10km between points, an effective 100-fold increase in detail from the original GCM resolution of 10000 km². The data were used to train a Deep Neural Network, which achieved an average prediction accuracy of $\pm 7\%$ from the observed rainfall value, with an R-squared value of 0.991. Analysis of the neural network suggests the significant importance of rainfall and humidity to model predictions, indicating that local relative rainfall distribution depends on rainfall and humidity rather than temperature. With this model, we enable local governmental and budgetary bodies to make informed decisions in response to future climate change impacts, such as for State Emergency services and Infrastructure development efforts.

KEYWORDS: Earth and Environmental Science, Water Science, Neural Network, Downscaling, Climate Model, GCM Downscaling.

■ Introduction

Australia is known for its natural wonders, but the same environment has been causing increasing damage due to climate change. According to the Insurance Council of Australia Insurance Catastrophe Resilience Report 2025, damages from climate-related extreme-weather events are projected to increase to \$35.2 billion in Australia by 2050, with climate change further increasing rainfall intensity and flood frequency in Australia's densely populated coastline. To reduce or prevent these potential damages, local and regional budgetary and planning bodies need access to useful data on future climate projections to make effective, informed decisions. Accurate future projections both allow first responders (State Emergency Services, Fire Service, Ambulance, Police) more time to prepare and provide present local governmental bodies with incentives to make sustainable decisions.¹⁻³

Climate models, specifically General Circulation Models (GCMs), have provided key insights into the potential global impacts of climate change scenarios, but they lack the high-resolution, local-scale detail of traditional weather forecasts for use by local government bodies.⁴ It is important to recognize the distinction between climate models and traditional weather forecasts: GCMs are primarily used to simulate future rainfall under defined atmospheric conditions,⁴ whereas weather forecasts predict near-future rainfall using concurrent observations of local climate.⁵ GCMs operate by physically simulating atmospheric circulation within a defined column of air above the ground, thereby requiring large simulation cells to capture complex large-scale atmospheric processes, but at the expense of output detail.⁴ To gauge climate change impacts,

climate model developers input a hypothetical climate scenario, for example, the atmosphere with 600 ppm carbon dioxide, into the GCM and analyze the output. For use in local and regional impact assessment, GCM output needs to be appropriately downscaled to a higher resolution.⁴

GCM downscaling is separated into two main methodologies: dynamical and statistical downscaling.⁶ Dynamical downscaling involves operating a climate model at a higher spatial resolution, producing output with higher spatial resolution.⁶ This can be done at a global scale or, more commonly, in Regional Climate Models (RCMs), which generate climate simulations within a defined area of the Earth. The downsides are that dynamical downscaling requires physical access to the climate model's code and programs to produce higher-resolution results, which is highly computationally intensive.^{2,6} It also relies on the creators of the GCM to manually produce downscaled results or to create a corresponding RCM, which means that GCM-derived RCMs for specific regions of the world, for example, the Australian state of New South Wales, are not available.³ Additionally, as this study did not develop or have the source code for a GCM, this study did not utilize the dynamical approach. Conversely, Statistical downscaling uses observed relationships between climate model outputs and local atmospheric conditions to train a separate model.⁶ Unlike dynamically downscaled models, statistical models utilize the GCM output only as an input variable, among other climatological variables, with a provided ground truth to compare the predicted rainfall of the model with actual rainfall.^{7,8} As such, statistical downscaling models do not require physical access to the GCM's operating code to make predictions, only its output

data, other gathered inputs, and observed variables.⁶ However, traditional methods of statistical downscaling generally require large, diverse datasets derived from complete atmospheric measurements from select weather stations,⁹ which are not available at a national scale. Moreover, they are often complicated models,^{7,8} requiring detailed, long-term observation of the ground truth for the model to construct a statistical relationship between the predictors and the observed values.⁶ As a result, traditional statistical downscaling models have the drawback of an overly narrow focus on local and catchment scales.^{1,9-11}

The method employed in this study (regressing rainfall using a deep learning neural network) is a statistical downscaling approach. This study seeks to build on existing literature by downscaling GCM outputs at the national scale, an area of orders of magnitude larger than those in previous Australian Literature, which often focuses on downscaling within a specific catchment basin.^{1,9-11} Research conducted in Australia, such as Timbal *et al.*, downscales GCM data within specific “areas of interest”, specifically being the Murray-Darling Basin, Nullarbor Plain, South East Coast, Tasmania, South East Australia, and South West Australia regions.¹ This reflects the aforementioned tendency for Australian climate downscaling studies to downscale within state and catchment-sized scales, not at the national scale. Internationally, studies such as Ahmed *et al.* and Polasky *et al.* downscale climate data at the state scale,^{9,11} with the former downscaling precipitation in the Balochistan region of Pakistan and the latter downscaling GCMs in the Midwest United States. To bridge this gap in national-scale GCM downscaling, a machine learning approach was used to achieve acceptable regression accuracy whilst simultaneously requiring less computational power and reducing the complexity of the required dataset. Previous studies that utilize a machine-learning downscaling approach include Ahmed *et al.* and Wilby *et al.* These studies gather large amounts of atmospheric and GCM data, as well as ground-truth data from measured, time-series climate data in the real world corresponding to the GCM variable.^{9,12} This study’s methodology is based on the methodologies established in the two ML-based papers above, outlined in Figure 1 with a flowchart.

■ Methods

Study area:

The study area covers the entire extent of continental Australia, including Tasmania. Australia features a predominantly arid interior, with temperate climates in the East.¹³ The Great Dividing Range in Australia’s east blocks rainfall from the sea from reaching the desertified interior, causing Australia’s east to experience temperate to tropical conditions with significant rainfall variability. In the states of Queensland and the Northern Territory, temperate grasslands gradually transition to tropical rainforest conditions, creating pockets of extremely high average annual rainfall.¹³

Data sources

The data used in this study were obtained from multiple sources. Decadal-averaged data and ground-truth annual observed rainfall were obtained from the Australian Bureau of Meteorology climate maps.¹⁴ The observed rainfall is available only at the monthly scale, so the extracted monthly values must be aggregated to obtain the 2024 yearly observed rainfall. To source elevation data, the GEODATA 9-second Digital Elevation Model 2008 3.0 was downloaded from Geoscience Australia,¹⁵ covering all of Australia. The climate model output used originated from the Italian CMCC-CM2-SR5 General Circulation Model,¹⁶ and was obtained via the European Network of Earth System Modeling (ENES) Climate4Impact portal.¹⁷ Specifically for this climate model, hindcast data for 2024 were used, which correspond to the observed 2024 rainfall data from the Australian Bureau of Meteorology, which served as the ground truth for the model.

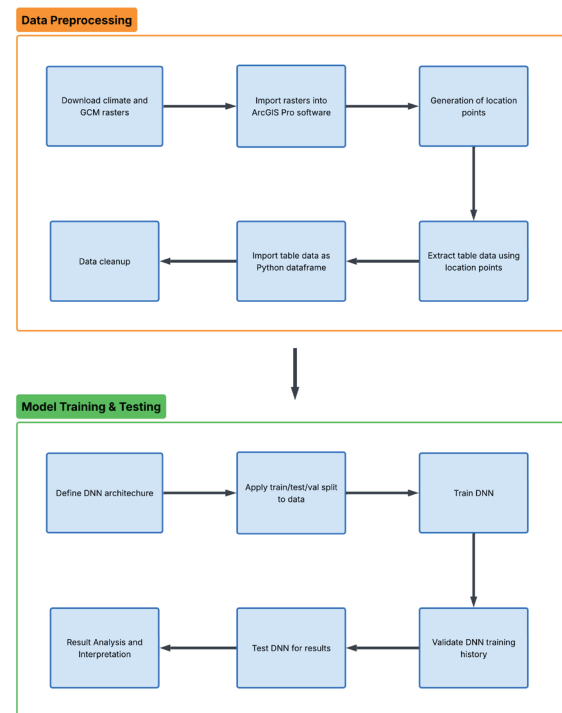


Figure 1: Methodology flowchart of this study. The methodology is split into two main stages in chronological order: the data collection and location and model training stages, respectively.

Data preprocessing:

The dataset was split into two components: predictors and labels. The predictors include spatially variable data or otherwise simulated data that lack a direct temporal reference. This includes historical averages of rainfall and humidity, derived from measurements over the past decades, and is assumed to be indicative of future observed values. Note that the GCM hindcast 2024 output is a predictor, as the GCM output is not dependent on observations made in the year 2024, but instead simulated using 2024’s expected atmospheric composition. Therefore, the GCM can generate outputs for future years and serve as a predictor. Other climatological variables were selected based on prior research, namely observed annual rainfall, annual average humidity (measured at 9 am and 3 pm),

elevation, and annual temperature (maximum, minimum, and mean).¹ As the model aims to replicate the long-term rainfall distribution within a large-scale GCM cell, an adjustment was made to use the spatially variable versions of these otherwise short observational time series. The rainfall, humidity, and temperature used are 30-year climatological averages across Australia. These values are assumed to be effectively constant when downscaling GCM outputs of near-future climate conditions. The label group comprises exclusively observed rainfall data, specified by the GCM hindcast date. As seen in Table 1 below, 8 input variables and 1 observed variable, against which the model output is compared, are used.

Table 1: Table of input variables used in the model with associated meaning and units. The 30-year climatology in this table is the same for all associated variables, being averages of the yearly values over the period 1990-2020.

Variable name	Meaning	Units
GCM_Hindcast_2024	Simulated rainfall value within a GCM cell, output as an average over a set year at a location	Millimeters (yearly total)
average_annual_rainfall	Total rainfall value averaged over a 30-year climatology (1990-2020) at a location	Millimeters (yearly total)
9am_humidity	Average 9 am humidity taken at a location over a 30-year climatology	% Relative humidity
3pm_humidity	Average 9 am humidity taken at a location over a 30-year climatology	% Relative humidity
elevation	Elevation above sea level at a location	Metres
mean_annual_temperature	The average of the mean yearly observed temperatures over a 30-year climatology	Degrees Celsius
min_annual_temperature	The average of the minimum yearly observed temperatures over a 30-year climatology	Degrees Celsius
max_annual_temperature	The average of the maximum yearly observed temperatures over a 30-year climatology	Degrees Celsius
2024_Yearly_Observed	Total observed rainfall at a location in the Year 2024.	Millimeters (yearly total)

a raster or a geographical heatmap. A raster map is composed of individual pixels of a specified geographical size resolution, and is combined to form a large map image of the specified raster area. Each pixel, also known as a raster cell, holds a specific value and is often visualized in Geographic Information Systems (GIS) software using a color gradient that denotes the relative magnitude of the value in the cell. As seen in Figure 2, these rasters can spatially overlap, necessitating their distinction using different color gradients.

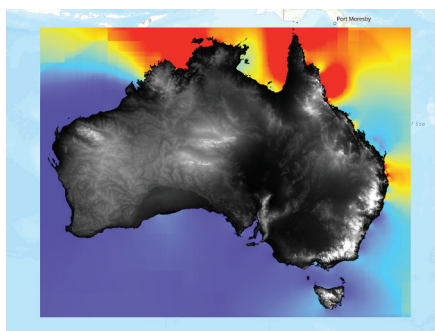


Figure 2: The elevation raster (black & white) overlaying the observed rainfall January 2024 raster (Colored). The color gradient in each raster (white to black and red to blue, respectively) indicates the relative magnitude of feature values in a raster cell. The rainfall raster lies beneath the elevation raster by virtue of the draw order in GIS software, demonstrating how the rasters used geographically overlap on the Australian Landmass.

Data preprocessing - Generation of datapoints:

Data in this study encompass a rectangular area with top-left and bottom-right bounds of 112.74°E, 10.27°S, and 156.04°E, 43.87°S, respectively. Because raw climatological data are typically available in raster form, they must undergo extensive preprocessing before they can be used in the neural network. The raster datasets were first loaded into ArcGIS Pro,¹⁸ resulting in multiple geographically overlapping raster datasets, as all rasters have defined data within Australia.

To convert this data into a table, coordinate points were generated in ArcGIS Pro for use with the Extract Multi Values to Points tool.¹⁸ The Extract Multi Values to Points tool in ArcGIS Pro operates by identifying the raster cell that spatially coincides with each input point and retrieving the cell value at that location.¹⁸ It performs this process for all selected raster layers (e.g., rainfall observed, GCM raster, humidity average). The result is a feature matrix in which each row corresponds to a point, and each column contains the extracted values from the corresponding raster datasets. As shown in Figure 3, the aforementioned coordinate points are individual locations. The exported feature matrix is ultimately used as the model's input, with each row corresponding to a location point and therefore a data sample.

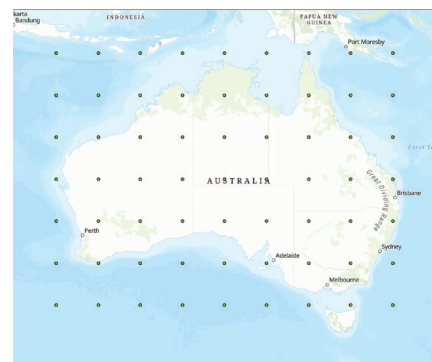
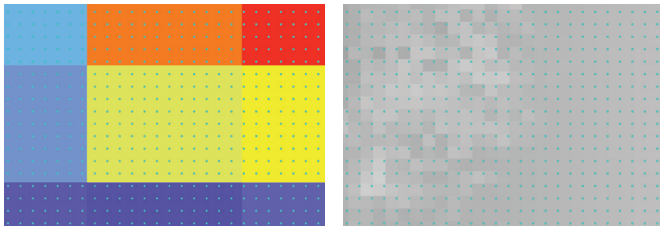


Figure 3: Simplified version of datapoints drawn in ArcGIS Pro.¹⁸ The actual data points used are far more numerous and distributed across a far denser grid than shown in the image, which occupies each humidity raster cell.

To downscale the GCM raster correctly, the point grid must be configured so that multiple points fall within each large GCM cell, while each point lies in a different cell for the higher-resolution predictor rasters. This ensures that only the climate model output is being spatially downsampled, while the other environmental variables contribute unique, location-specific information to generate valid predictions. To achieve this, the point density was set to the cell size of the lowest-resolution predictor raster beside the GCM raster - the humidity raster with a resolution of $0.1^\circ \times 0.1^\circ$ (approximately 10 km $\times$ 10 km). Using this spacing, 146,258 coordinate points were generated, forming a continuous point grid across the entire contiguous Australian mainland and Tasmania.

Figure 4 below illustrates the concentration of data points by overlaying the data point grid on the GCM and humidity rasters in a given region. Specifically, the figure overlays the Australian coastal city of Port Macquarie, which was arbitrarily chosen for this example. As seen in Figure 4(a), 96 data points are located within the central GCM raster cell depict-

ed in yellow, which downscales the GCM raster and predicts rainfall values for each data point. As seen in Figure 4(b), each humidity raster cell contains only one data point, providing the model with a spatially unique humidity value for a valid prediction. Thus, Figure 4(b) shows the maximum spatial density of data points possible for this study, which was ultimately used for data extraction. It is important to note that a less spatially dense data point grid can be used for data extraction. A valid prediction does not require that each humidity raster cell be used for all input data, just that the data points must each represent spatially unique data, where no two data points can occupy the same cell of any non-GCM rasters. In higher-resolution rainfall rasters, for example, the temperature raster, the data points occupied only one in four cells on average.



(a) GCM Raster. The central grid cell size is 100 km by 100 km. Surrounding cells are cropped. The color gradient from red to blue represents high to low-value cell attribute values, respectively. Blue dots are data points spaced 10 km by 10 km. (b) Humidity Raster. Cell size is 10 km by 10 km. Edge cells are cropped. The color gradient from white to black represents high to low cell attribute values, respectively. Blue dots are data points spaced 10 km by 10 km.

Figure 4: Generated data points (blue) overlaying the GCM Raster (a) and the Humidity raster (b). Around 100 data points are contained within each GCM raster cell, while a humidity raster cell only contains 1 data point. The area is an aerial shot of Port Macquarie, New South Wales, Australia.

Because some original data point locations contained many values over the ocean, the data needed to be filtered to eliminate data points that lacked complete data across all selected rasters. This is because the individual rasters have different sizes: the GCM raster is global-scale, the elevation raster covers only Australian land, and others cover land and the nearshore ocean. The simplest way to filter is to delete all data from datapoints without a specified elevation value, as the elevation raster is the smallest and is included within all other raster maps, meaning that a datapoint geographically overlapping the elevation raster must therefore be included in all other raster maps as well. This leaves 69581 usable datapoints, from the original 146258, for training and testing the model.

All monthly observed rainfall totals were summed to obtain the 2024_Yearly_Observed value, which was used as the label. The GCM_Hindcast_2024 raster data was averaged over a year to mm/s, then multiplied by 31556926 (the number of seconds in a year) to extrapolate the yearly GCM value for 2024, ensuring all rainfall variables were in units of mm/year.

Deep Neural Network (DNN) Model:

A deep neural network was used in this study, based on prior findings that neural networks consistently yielded lower Root Mean Squared Error (RMSE) values than traditional statistical downscaling methods, which instead used first-order two-state Markov processes, binomial occurrences, and vor-

ticity-based resampling.¹² Although DNNs noticeably suffer when attempting to predict the random occurrence of rainy days,¹² this study does not face this limitation. As this study performs downscaling at the annual time scale, each prediction represents the sum of many daily rainfall values. When many daily observations are aggregated, the resulting annual total tends to exhibit a more symmetric and stable distribution. Under the Central Limit Theorem, the sum of a large number of daily rainfall amounts with finite variance will approximate a normal distribution, reducing sensitivity to inaccuracies in simulating individual extreme wet days.

Model Architecture:

The DNN has 8 input neurons, 6 hidden layers with 30 neurons each, and 1 output neuron. This specific combination of parameters was selected through a manual search in hyperparameter tuning, which aims to identify the best-performing model architecture. This model is a Deep Neural Network (DNN) because it has more than one hidden layer. The Rectified Linear Unit (ReLU) activation function was applied between layers, constraining values to either zero or positive.

The 8 input neurons and 1 output neuron are defined by the number of selected input variables and the required regression outputs, which are fixed and cannot be changed. Therefore, tests to find the optimal model architecture were conducted using a grid search to determine the number of hidden layers and the number of neurons per hidden layer. All tested model parameters are shown below, in increments of 1 for hidden layers and 10 hidden neurons for a total of 25 manual tests.

Tested number of hidden layers: [3, 4, 5, 6, 7]

Tested hidden neurons per hidden layer: [10, 20, 30, 40, 50]

DNN Training Loop:

The training loop is configured to take a data batch size of 10, and Mean Squared Error (MSE) loss was used in the model. This process includes computing the error as a function of the model weights, then finding the gradient as the partial derivative of the loss with respect to the model weights. The model weights are then updated in the opposite direction, reducing the error in a process known as gradient descent. In this study, the Adam optimizer was used.¹⁹

The overall dataset was randomly split into training (64%, 44531 samples), validation (16%, 11133 samples), and test (20%, 13917 samples) sets. The model was set to train for 1500 epochs. To ensure model convergence and prevent overfitting, early stopping was implemented with a patience of 100 epochs.

Results and Discussion

The model with 6 Layers and 30 hidden neurons per layer achieved the lowest Root Mean Square Error (RMSE), while simpler networks failed to capture and predict rainfall values with acceptable accuracy. Models with more hidden layers and more hidden neurons achieved similar accuracy but required significantly longer training times and were therefore deemed suboptimal.

The model converged after 1338 epochs of training, with the MSE reduction following an exponential decay pattern, as

shown in Figure 5. The neural network's predictions on the test set had an RMSE of around 40mm/year, with the average label value being around 600mm/year. This represents an effective error of $\pm 7\%$ relative to the predicted value, which is within the acceptable $\pm 25\%$ accuracy threshold for short-term rainfall forecasts.⁵

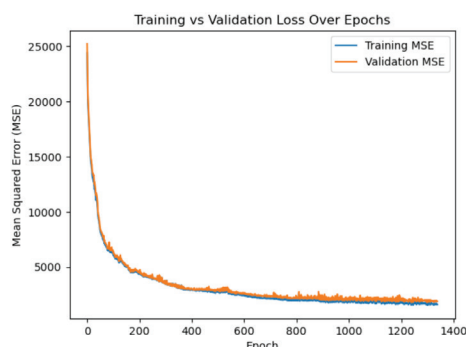


Figure 5: Model error across epochs of training. MSE on the training and validation sets was graphed separately to inspect the model for overfitting, which would cause Validation MSE to be significantly higher than Training MSE. The close alignment of training and validation MSE curves demonstrates that the model has not overfitted to the training data.

An R-squared value indicates the percentage of variance in the predicted values explained by the variance among the predictor variables. In simple terms, it is the percentage variation in the final value that can be explained by the predictors. A high R-squared value of 0.9–1 indicates that the predictors account for the majority of the variation in the output variables, indicating a statistically valid prediction. R-squared alone, however, does not reveal the full extent of a model's accuracy, as systematic biases towards certain predictor variables or value ranges can reduce the model's prediction quality while maintaining a high R-squared value. Therefore, this study uses additional methods, such as RMSE, predicted vs. actual, and residual vs. actual graphs, as well as SHAP analyses, to fully assess the model's accuracy across all predicted and error ranges. The results of those tests are outlined further below.

The R-squared value calculated between the observed and predicted values of the model averaged around 0.991, meaning that 99.1% of the variance in the observed values can be explained by the variance in the predicted values. This suggests that the model and its input variables account for the vast majority of factors contributing to the final yearly total rainfall, reinforcing the $\pm 7\%$ accuracy previously found with RMSE. R-squared values this high may suggest data leakage between the predictor variables and the target label, undermining the validity of the model's predictions. In this case, the main concern is potential leakage between the GCM hindcast rainfall predictor and the observed ground-truth rainfall used as the label. This is because observed rainfall totals are incorporated into the GCM hindcast dataset to bias-correct the model output toward real observations when generating the hindcast rasters of the CMCC-CM2-SR5 General Circulation Model.¹⁶ However, this hindcast data is generated at the decadal scale,¹⁶ averaged over the years 2014–2024. The data used to validate the climate model hindcast must therefore have been captured in 2014, which validates the R-squared value, as

the possibility of the GCM bias-correcting to the 2024 annual observed rainfall value was ruled out. Furthermore, the GCM data was installed before the migration of the European Network for Earth Systems Modeling (ENES) server infrastructure in December 2024, ruling out data leakage, as the Australia-wide annual rainfall total data for 2024 did not yet exist.

As seen in Figure 6 below, the model predictions closely align with the theoretical perfect accuracy, with RMSE = 0, as denoted by the line $y = x$ across the entire range of predicted and actual values. Although most values cluster in the 0–2000 mm rainfall per year range in Figure 6, the extremes in the 3000–5000 mm range also align closely with the line, suggesting that the model provides accurate predictions for both low and extremely high rainfall in Australia. This also suggests that the earlier RMSE value is not distorted by any significant outliers, and that an error within $\pm 7\%$ of the observed value is an effective measure of model accuracy.

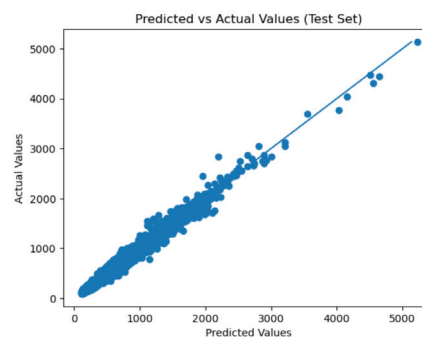


Figure 6: Predicted vs Actual values graph. The predicted vs actual values graph is used to gauge model performance across a wide range of values. Each dot in the graph represents a single prediction from the model on the test split of the dataset. This graph pattern indicates that the model has low error across the entire range of predicted rainfall values.

As illustrated in Figure 7, the distribution of residuals against actual values shows no discernible patterns or systematic biases. This uniform performance validates the high R-squared and RMSE values, ensuring they are not skewed by specific data ranges. The clustering of the dots in the 0–2000 mm range on the left side of Figure 7 indicates only the frequency of data points with actual values in that range and does not indicate any systematic error.



Figure 7: Residuals vs actual scatter plot. Each dot is a model prediction on the test set, with the 0 line indicating a perfect prediction. Dots above and below the line represent overpredictions and underpredictions, respectively. Randomly scattered dots above indicate that the model is not systematically biased towards any range.

A SHAP analysis was used to quantify the individual contributions of selected variables to the neural network's output.²⁰ As seen in Figure 8, the variable with the greatest contribution is the decadal average annual rainfall variable, due to its greatest range of SHAP values, followed by the GCM output, and then closely by the two humidity variables. This analysis suggests that the model is utilizing predominantly rainfall and humidity values for prediction, indicating that local relative rainfall distribution is determined predominantly by rainfall and humidity variations rather than temperature.

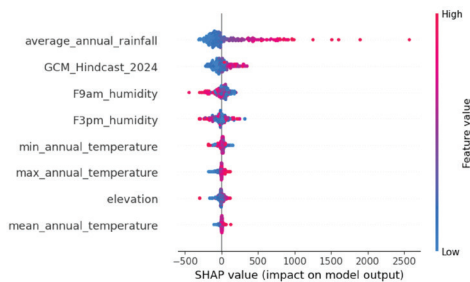


Figure 8: SHAP analysis of the neural network. The 8 input variables are listed on the left, with their contributions shown on the bottom as a scale. Each point in the middle is an instance of the corresponding input variable, with the point's SHAP value coming from a single prediction of the test set. This allows us to interpret the DNN and gauge which climate phenomena contribute to annual rainfall the most.

Limitations and considerations:

The model, by making a prediction, effectively simulates rainfall distribution within a specified General Circulation Model Cell at a 10 km by 10 km resolution. This study does not use time-series data as predictors. The model views the modeled rainfall distribution as a static factor that dictates predictions. This distribution was modeled based on the 30-year average annual climatology for 1990–2020, meaning the predictions could be reasonably extrapolated into the near future. An obvious limitation to this approach is that climate conditions are not static. This means that future changes in climate composition and atmospheric processes due to climate change cannot be accurately modeled using the 30-year climatological average data used in this study. Therefore, the model needs to be periodically retrained using new decadal-average rainfall data.

This model also assumes that the relative rainfall distribution within the spatial extent of a GCM cell will remain constant across different climate conditions. This assumption was made because climate change alters predominantly mean precipitation levels across a wide region rather than the relative distribution or concentration of precipitation within a select area,²¹ suggesting that regional relative rainfall distribution is a product of temporally static, likely topological factors. Without this assumption of a static, localized, relative rainfall distribution, the model would need to physically simulate rainfall and precipitation processes in a 100 km by 100 km area for each data point, which is not within the scope of this study, as it would effectively make the model a dynamically downscaled GCM.

To explain this assumption, imagine an arbitrary GCM grid cell that overlaps City A and City B, with City A receiv-

ing twice the amount of rainfall as City B. In a hypothetical climate scenario, this study assumes the relative rainfall distribution within a GCM cell is constant, meaning that City A still receives twice the rainfall as City B despite drastic changes in the area-average rainfall. While near-future local rainfall distribution could be modeled statically under this assumption, changes to anthropogenic land use, significant changes to regional topography, and significant changes to climate conditions could alter this rainfall distribution, rendering model predictions for the far future or under extreme climate conditions invalid with high error.

Future directions:

While this study effectively predicts annual rainfall at a specified location, it does not estimate the frequency or temporal distribution of rainfall within a specific year. Future research could use the initially separate monthly ground-truth values to predict rainfall at the monthly temporal scale, or use daily total data from the Australian Bureau of Meteorology to predict rainfall at the daily temporal scale. Alternatively, future research could independently model the temporal distribution of rainfall over a year, using the data generated by this study's model as an input variable. Physical climate simulations could also be used in the aforementioned research, given the El Niño Southern Oscillation's significant impact on Australian precipitation distribution,¹³ among other factors not included in this study, such as wind speed, air pressure at set altitudes, and surface divergence.^{9,22}

Future research could also aim to increase the resolution of the downscaled data by using a higher density of data points than in this study. This study was limited by the coarse resolution of the 9 am and 3 pm humidity datasets, which necessitated downsampling to 10 km by 10 km. Although useful for regional and local-scale applications, this data could be considered too coarse for extremely narrow use cases, such as suburb-scale planning and budgetary concerns. As such, research utilizing a finer-resolution humidity raster at the national scale can significantly increase detail, further expanding the study's possible applications. A possible barrier is that high-resolution climatological raster datasets are rare at the national scale, which could be addressed by splicing multiple high-resolution local rasters into a national raster for a specified variable. For specialized regional applications of GCMs that require higher resolution, this machine learning approach could be used in conjunction with dynamically downscaled RCM rasters for a defined region to achieve even higher-resolution and more accurate predictions. As dynamically downscaled RCMs typically have a resolution of 1–2 km per grid cell, a further statistical downscaling model could result in predictor data point separation of 100–200 m, a 10-fold increase in downscaled resolution from this study.

Conclusion

In this study, we downscaled the output of the CMCC-CM2-SR5 General Circulation Model at a national scale in Australia, the largest scale of climate model downscaling

achieved in Australian literature. The DNN model used in this research predicts rainfall with an average prediction accuracy of $\pm 7\%$ relative to the mean and an R-squared value of 0.991, well within the acceptable error bound of $\pm 25\%$ for weather forecasting. The SHAP analysis suggests a greater contribution of the GCM, rainfall, and humidity variables than of temperature, indicating that local relative rainfall distribution is determined predominantly by rainfall and humidity variations rather than temperature. The downscaled resolution of 10 km by 10 km, effectively a 100x increase in detail from the original climate model output, allows local councils and budgetary bodies to make informed decisions in response to future climate change. Through this study, we lay a stepping stone in the nationwide availability of GCM model results, which we hope will allow future researchers to explore this critical field further.

■ Acknowledgments

This study would like to thank the Currie Country Social Change organization for their continued support of our research, informing our previous studies as well as similar flood research studies to the Australian Parliament regarding flood impacts on Indigenous Australians.

■ References

- Timbal, B., Fernandez, E., & Li, Z. (2009). Generalization of a statistical downscaling model to provide local climate change projections for Australia. *Environmental Modeling & Software*, 24(3), 341–358. <https://doi.org/10.1016/j.envsoft.2008.07.007>
- Xu, Z., Han, Y., & Yang, Z.-L. (2018). Dynamical downscaling of regional climate: A review of methods and limitations. *Science China Earth Sciences*, 62. <https://doi.org/10.1007/s11430-018-9261-5>
- Lpez-Gmez, I., Wan, Z. Y., Zepeda-N. ez, L., Schneider, T., Anderson, J., & Sha, F. (2024). *Dynamical-generative downscaling of climate model ensembles* (arXiv:2410.01776). arXiv. <https://doi.org/10.48550/arXiv.2410.01776>
- Intergovernmental Panel On Climate Change (IPCC). (2023). *Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (1st ed.). Cambridge University Press. <https://doi.org/10.1017/9781009157896>
- Shahrban, M., Walker, J. P., Wang, Q. J., Seed, A., & Steinle, P. (2016). An evaluation of numerical weather prediction-based rainfall forecasts. *Hydrological Sciences Journal*, 61(15), 2704–2717. <https://doi.org/10.1080/02626667.2016.1170131>
- Tang, J., Niu, X., Wang, S., Gao, H., Wang, X., & Wu, J. (2016). Statistical downscaling and dynamical downscaling of regional climate in China: Present climate evaluations and future climate projections. *Journal of Geophysical Research: Atmospheres*, 121(5), 2110–2129. <https://doi.org/10.1002/2015JD023977>
- Sachindra, D. A., Huang, F., Barton, A., & Perera, B. J. C. (2014a). Statistical downscaling of general circulation model outputs to precipitation—part 1: Calibration and validation. *International Journal of Climatology*, 34(11), 3264–3281. <https://doi.org/10.1002/joc.3914>
- Sachindra, D. A., Huang, F., Barton, A., & Perera, B. J. C. (2014b). Statistical downscaling of general circulation model outputs to precipitation—part 2: Bias-correction and future projections. *International Journal of Climatology*, 34(11), 3282–3303. <https://doi.org/10.1002/joc.3915>
- Ahmed, K., Shahid, S., Harun, S., & Wang, X. (2015). Multilayer perceptron neural network for downscaling rainfall in arid region: A case study of Baluchistan, Pakistan. *Journal of Earth System Science*, 124, pp. 1325–1341. <https://doi.org/10.1007/s12040-015-0602-9>
- Liang, X.-Z., Kunkel, K. E., Meehl, G. A., Jones, R. G., & Wang, J. X. L. (2008). Regional climate models downscaling analysis of general circulation models presents climate biases propagation into future change projections. *Geophysical Research Letters*, 35(8). <https://doi.org/10.1029/2007GL032849>
- Polasky, A. D., Evans, J. L., Fuentes, J. D., & Hamilton, H. L. (2022). Statistical climate model downscaling for impact projections in the Midwest United States. *International Journal of Climatology*, 42(5), 3038–3055. <https://doi.org/10.1002/joc.7406>
- Wilby, R. L., Wigley, T. M. L., Conway, D., Jones, P. D., Hewitson, B. C., Main, J., & Wilks, D. S. (1998). Statistical downscaling of general circulation model output: A comparison of methods. *Water Resources Research*, 34(11), 2995–3008. <https://doi.org/10.1029/98WR02577>
- Raktima, D., Lewis, S. C., Alblaster, J. M., Abram, N. J. (2019). A review of past and projected changes in Australia's rainfall. *Wires Climate Change*. <https://doi.org/10.1002/wcc.577>
- Australian Government Bureau of Meteorology. Climate. <https://www.bom.gov.au/climate/> (accessed 13 June 2024)
- Australian Government Geoscience Australia. Digital Elevation Data. <https://www.ga.gov.au/scientific-topics/national-location-information/digital-elevation-data> (accessed 23 March, 2025)
- Nicoli, D., Bellucci, A., Ruggieri, P., Athanasiadis, P., Materia, S., Peano, D., Fedele, G., & Gualdi, S. (2022). *The CMCC Decadal Prediction System*. Climate and Earth system modeling. <https://doi.org/10.5194/gmd-2022-181>
- IS-ENES. Climate4Impact. <https://www.climate4impact.eu/c4i-frontend> (Accessed November 19, 2024)
- Esri. ArcGIS Pro, version 2025 3.0.1; Environmental Systems Research Institute: Redlands, CA, 2025.
- Kingma, D. P., & Lei, J. (2015). *Adam: A Method for Stochastic Optimization*.
- Lundberg, S., & Lee, S.-I. (2017). *A Unified Approach to Interpreting Model Predictions* (arXiv:1705.07874). arXiv. <https://doi.org/10.48550/arXiv.1705.07874>
- Nazeri Tahroudi, M. (2025). Comprehensive global assessment of precipitation trend and pattern variability considering their distribution dynamics. *Scientific Reports*, 15(1), 22458. <https://doi.org/10.1038/s41598-025-06050-5>
- Rampal, N., Gibson, P. B., Sood, A., Stuart, S., Fauchereau, N. C., Brandolino, C., Noll, B., & Meyers, T. (2022). High-resolution downscaling with interpretable deep learning: Rainfall extremes over New Zealand. *Weather and Climate Extremes*, 38, 100525. <https://doi.org/10.1016/j.wace.2022.100525>

■ Author

Alex Huijie Zhuang studies at The King's School, Parramatta. Alex is a former member of the Australian National Earth Science Olympiad Team, winning a Gold individual medal at the 2025 International Earth Science Olympiad. His research in flood prediction was recognized at ISEF 2024 with a Third Place Award. Contact at zhuanghuijie2008@gmail.com

Machine Learning in Predicting Outcomes of Brain Tumor Resection Surgeries: A Scholarly Review

Arman Raoufian

Northwest High School, 13501 Richter Farm Rd, Germantown, Maryland, 20874, USA; infectlate@gmail.com

ABSTRACT: Being able to predict the outcome of a patient following a brain tumor removal is essential to providing care for those patients. Conventional tools for predicting the outcome of patients who undergo a brain tumor removal are not able to accurately model the complex relationships between the various variables of each patient and their condition. Machine learning, however, has the potential to provide a more powerful framework for predicting the survival of patients after brain tumor removal, the complications that may occur following those operations, and the quality of life of those patients following the operations. Through searching databases like PubMed and Google Scholar for articles published between 2015 and 2025, it becomes evident that based models have higher accuracy in predicting the outcomes of patients after brain tumor operations than conventional models. One systematic review of studies that utilized machine learning to model the outcomes of patients after brain tumor operations discovered that the median area under the curve (AUC) score of the models was 0.83, with a median accuracy of 94.5%. Despite these positive findings, however, various barriers to the adoption of these models into the clinical care of brain tumor patients exist. As a result, models will likely be utilized as a supplement to the clinician's care for those patients rather than as a means of replacing the clinician's judgments of those patients' outcomes.

KEYWORDS: Biomedical Engineering, Biomedical Imaging, Machine Learning, Brain Tumor Resection, Outcome Prediction.

■ Introduction

Brain tumor resection aims to remove as much of the tumor mass as possible while preserving critical functions of the brain, such as speech, movement, and memory. The recovery, survival, and quality of life of patients who undergo brain tumor resection depend on various factors, such as the location of the tumor within the brain, the grade of the tumor's histology, and the age and general health of the patients.¹ Each of these factors can have a substantial effect on the outcome of the resection procedure. Accordingly, it is critical to be able to accurately predict the outcome of brain tumor resection for each patient.

Various scoring systems and statistical models are used to attempt to estimate the outcome of brain tumor resections. These models are helpful in providing an estimate of the outcome of brain tumor resection for each patient. However, these models often evaluate each of the model variables separately and make assumptions regarding the linear relationships between those variables and the outcome of the brain resection procedure.² As a result, these models may not account for the effects of nonlinear relationships between the variables on the outcome of the resection of brain tumors. Therefore, there is a need for improved tools for predicting the outcome of resection of brain tumors.

Machine learning models can help address the shortcomings of conventional models for predicting the outcome of brain tumor resection. Machine learning models learn the relationships between various variables from data without having to prescribe a specific model form to the data.³ By training a machine learning model on historical records of brain tumor

resections, machine learning models can provide individualized predictions regarding the outcome of a brain tumor resection procedure for a newly admitted patient. In this review, the literature on the use of machine learning models for predicting the outcome of brain tumor resection is summarized and evaluated for its ability to produce models that can more accurately predict the outcome of the procedure than the conventional models for predicting its outcome.

■ Methodology

This paper is a narrative literature review with a focus on answering the question of to what extent does machine learning, as it pertains to the field of neurosurgery and brain tumor resection surgery, provides stronger predictions of patient outcomes compared to the traditional models that have been used in neurosurgery to date.

A search was performed on the databases PubMed and Google Scholar, utilizing the keywords: machine learning, brain tumor resection, neurosurgery outcome prediction, survival modeling, and postoperative complications. The search was limited to peer-reviewed journals in the English language between January 2015 and April 2025.

The studies were included if they utilized machine learning to make predictions of patient outcomes in the field of neurosurgery, reported some form of performance metrics for their models, and their outcomes related to the resection of brain tumors. Studies were excluded if they did not report performance metrics for their models, they were not related to neurosurgery, they were not peer reviewed, or they were conference abstracts without a published article with the same content.

There were over 200 articles found during the initial database search. However, only 22 fulfilled the requirements for inclusion within the review. These in-depth studies will be summarized in this literature review according to the following main topics: traditional models, machine learning models, evidence, and challenges. The findings presented in this review are a synthesis of the current literature on the topic, rather than a systematic review of the literature in the way that would be required for a PRISMA study.

Background: Traditional Approaches to Prognostication in Neurosurgery:

Prognostication is the estimation of the likely outcome of a patient's current medical condition. Because the conditions treated by neurosurgeons are often life-threatening conditions, it is essential to be able to accurately estimate the likely outcome of a patient with such a condition.⁴

Clinical Scoring Systems:

Experienced clinicians formulate an impression of the risk based on the history, physical examination, physiological, and intraoperative findings. While this method incorporates the clinician's judgment at the time of the procedure, the risk of inter-observer variability and bias makes this method less suitable for comparing the risk between institutions.⁴

Clinical Scoring Systems:

Several standardized scales have been established to assess the prognosis of TBI. The Glasgow Coma Scale (GCS) evaluates the level of consciousness of the patient and is the most commonly used scale in the evaluation of traumatic brain injuries.⁵ The Hunt and Hess scale is used to assess the neurological deficits of patients who suffer from subarachnoid hemorrhage.⁶ The Karnofsky Performance Scale Index evaluates the functional status of cancer patients and is inversely correlated with the mortality rate of those patients.⁷ These scales allow for better communication between healthcare team members; however, they all have the disadvantage of reducing the complex clinical presentation of a patient to a series of numbers, they can only assess certain static variables of the patient, and they do not consider the interaction between the various variables that contribute to the patient's outcome.⁸

Able to carry on normal activity and to work; no special care needed	100	Normal no complaints; no evidence of disease
	90	Able to carry on normal activity; minor signs or symptoms of disease
	80	Normal activity with effort; some signs or symptoms of disease
Unable to work; able to live at home and care for most personal needs; varying amounts of assistance needed.	70	Cares for self; unable to carry on normal activity or to do active work.
	60	Requires occasional assistance, but is able to care for most of his personal needs.
	50	Requires considerable assistance and frequent medical care.

Unable to care for self; requires equivalent of institutional or hospital care; disease may be progressing rapidly.	40	Disable; requires special care and assistance.
	30	Severely disabled; hospital admission is indicated although death not imminent.
	20	Very sick; hospital admission necessary; active supportive treatment necessary.
	10	Moribund; fatal processes progressing rapidly.
	0	Dead

Figure 1: Karnofsky Performance Scale. The Karnofsky Performance Scale is used to assess a patient's functional status and ability to perform daily activities. Scores range from 100 (normal, no complaints) to 0 (death), with lower scores indicating decreased functional ability and poorer clinical outcomes. This scale is commonly used in neurosurgery to help evaluate prognosis and guide treatment decisions. Created by Raoufian using Canva.

Imaging-based Prognostication:

Neuroimaging allows for the collection of objective and reproducible prognostic data. An MRI can reveal data about the enhancement of the tumor, the presence of edema, diffusion restriction, and mass effect, all of which are correlated with the grade of the tumor and the difficulty that may be encountered during surgical removal of the tumor.⁹ However, interpretation of the MRI reports relies upon the expertise of the interpreting radiologist to recognize patterns of the tumor on the MRI, and any subtle findings that are not visible to the unaided eye can be easily missed.¹⁰

Annotated Brain MRI Scans Highlighting Tumor Regions for Predictive Modeling

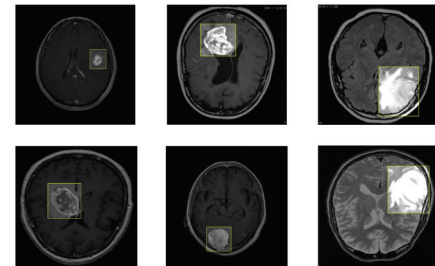


Figure 2: Axial brain MRI images that display the tumor regions outlined by bounding boxes and utilized as the inputs for the machine learning models. The variety in the size, shape, intensity, and location of the tumors in the brain illustrates the heterogeneity of these types of tumors. Each of these features can be extracted from the MRI images to analyze the tumors and predict the outcome of the surgical procedures performed on patients with these types of tumors. Source: Roboflow Brain MRI Dataset (<https://universe.roboflow.com/brain-mri/mri-rskcu>). No changes were made to the data. (CC BY 4.0).

Thus, standardized, objective, and reproducible, these results, too, are confirmed by the professional as there are certain atypical findings and those visible to the naked eye.¹⁰

Machine Learning Methods in Neurosurgical Outcome Prediction

Machine learning encompasses a variety of algorithms that computers use to learn from examples in the training data instead of relying on predetermined equations to calculate the outcome.¹¹ Medical records of patients who underwent certain neurosurgical procedures in the past are used to train machine learning algorithms to predict the outcome that will be observed for new patients undergoing those same procedures.² Because machine learning algorithms can examine numerous variables simultaneously to find correlations between variables that impact the outcome of surgical procedures, they have

attracted the interest of researchers who wish to incorporate these methods alongside the tools that physicians already utilize today.

Supervised Learning Methods:

The most common form of supervised learning is logistic regression. Using logistic regression, the probability that a patient will have a positive or negative outcome from a surgical procedure is calculated from various features of the patient and the procedure. The method is simple to implement, but it cannot account for the effects of nonlinear relationships between variables.² Decision trees are another form of supervised learning. Decision trees model the outcome of a surgical procedure by assigning thresholds to various features of the patient or the procedure. Each internal node within a decision tree represents some feature of the patient or procedure, and each branch descending from that node contains the possible values for that feature.¹³ Using only single decision trees to model surgical outcomes leads to overfitting, which introduces other methods into the equation.

Support vector machines (SVMs) find the optimal boundary that separates examples in training datasets with different outcomes. SVMs can model nonlinear relationships between features and surgical outcomes.¹⁴ As such, they can offer better discrimination between outcomes than logistic regression, but at the cost of interpretability.

Ensemble Methods:

Ensemble methods such as random forests and gradient boosting algorithms create groups of decision trees and combine the results from each individual tree within the group to find the outcome of the surgical procedure for a patient. Random forests use a group of decision trees that are built independently of each other.¹⁵ Gradient boosting algorithms build decision trees sequentially such that each tree corrects the errors made by the previous tree within the ensemble. The most common form of gradient boosting algorithms is gradient boosting and has been used in a variety of neurosurgical outcome prediction studies.¹⁹

Deep Learning:

Deep learning methods use convolutional neural networks (CNNs) to calculate the outcome of neurosurgical procedures. CNNs can analyze medical images of the patients to determine features related to the outcome of the procedure without having to create features manually.¹⁶ Deep learning methods achieve the highest accuracy with large datasets of medical records and can simultaneously process various types of data, from medical images to medical records.¹⁷

Evidence from Studies

The most comprehensive review of the use of ML models to predict neurosurgical outcomes was that by Senders *et al.* in 2018.² Senders *et al.* reviewed 30 studies using ML methods for outcome prediction after various types of neurosurgical procedures. Across the reviewed studies, ML models

achieved median accuracy rates of 94.5% and AUCs of 0.83 in predicting outcomes after procedures. Furthermore, ML models achieved a median 15% improvement in accuracy and 0.06 improvement in AUC when compared to models based on logistic regression. Additionally, in some studies, ML models outperformed both established outcome indices and expert surgeons in predicting outcomes of patients undergoing procedures. Thus, these results indicate that the advantages of ML models over conventional statistical models for outcome prediction in neurosurgery are consistent across multiple studies.

In the specific domain of outcome prediction following brain tumor resection, van Niftrik *et al.* performed studies comparing various boosting algorithms to conventional statistical models in predicting outcomes in 668 patients undergoing elective intracranial tumor surgeries at a single center.¹⁹ The outcomes that were assessed included the development of complications within the first 24 hours of surgery. Gradient boosting algorithms were the performing models in this group of studies, achieving an accuracy of 0.70 and an AUC of 0.73, while the conventional statistical model achieved accuracy rates of 0.59 and an AUC of 0.64. However, these studies were limited to only one surgical center.

Farrokhi *et al.* performed studies that applied supervised machine learning algorithms to predict outcomes following brain stimulation surgeries in 501 patients from multiple surgical centers. Models were developed using 70% of the patients' data and tested on the remaining 30%.²⁰ Gradient boosting machines were among the performing supervised machine learning algorithms for outcome prediction after brain stimulation surgeries. For example, machine learning algorithms achieved AUC values of 0.86 for any surgical complications, 0.91 for complications within the first 12 months after surgery, 0.88 for the need to return to the operating room, and 0.97 for infection after surgery. Factors that were found to impact the outcomes of deep brain stimulation surgeries included the age of the patient, their body mass index, the side on which the surgery was performed, the sex of the patient, and whether they had Parkinson's disease.

In efforts to use MRI scan images to determine the survival of patients with brain tumors, Chato and Latifi assessed the performance of various machine learning algorithms on the BraTS 2017 data set of 163 patients with gliomas and associated MRI scan images and survival times.¹⁶ The features that were used to train the machine learning models included volumetric, statistical, intensity texture, histogram, and deep learning features of the patients' MRI scans. The best results were achieved with machine learning models that used deep learning features of the patients' MRI scans, indicating that features extracted from MRI scans by deep learning methods outperform conventional features used by conventional machine learning classifiers.

In addition to the use of machine learning models trained directly with electronic health record data, Ive *et al.* performed a study in which natural language processing (NLP) models were used to extract information from the electronic health records of 2,268 patients undergoing elective neurosurgical procedures.³ A Random Forest machine learning model that

utilized information extracted by natural language processing models achieved a recall of 0.87 for identifying patients who would experience unplanned admissions to the intensive therapy unit (ITU), with an AUC of 0.99. Furthermore, by using this model, the proportion of patients with unplanned ITU admissions that were missed by the attending neurosurgery experts was reduced from 36% of cases to just 4%.

Badjatia *et al.* performed a review of the literature on the use of machine learning models to predict outcomes following traumatic brain injuries (TBIs).¹² The authors found that machine learning models that incorporated various types of data from the patients following TBIs, including clinical features, neuroimaging data, and autonomic nervous system metrics, can improve the prediction of outcomes following TBIs relative to conventional methods of predicting those outcomes. The review also indicated that many current machine learning models have challenges with issues like data variability, model interpretability, and overfitting. Thus, the authors call for the standardization of these models and the testing of these standardized models outside of the groups with whom they were developed.

Table 1: Summary of Primary Studies: ML Models for Neurosurgical Outcome Prediction

Study (Author, Year)	ML Model(s) Used	Dataset / Sample	Outcome Predicted	Performance Metrics	Comparison Baseline
Senders <i>et al.</i> , 2018	Systematic review of multiple ML model types (LR, ANN, SVM, RF, gradient boosting)	Review of 30 studies (PubMed/Embase search to Jan 2017)	Survival, recurrence, symptom improvement, adverse events across neurosurgical conditions	Median accuracy 94.5%; median AUC 0.83; ML outperformed LR by median 15% accuracy and 0.06 AUC	Logistic regression; established prognostic indices; clinical experts
van Niftrik <i>et al.</i> , 2019	Gradient boosting algorithms (multiple compared); best model selected from out-of-sample test set	668 patients; single-center prospective registry (University Hospital Zürich, June 2015–May 2017)	Early postoperative complications within 24 h of intracranial tumor surgery (Clavien-Dindo classification)	Accuracy 0.70 (CI 0.59–0.79); AUC 0.73; sensitivity 0.80; specificity 0.67. Conventional model: accuracy 0.59, AUC 0.64	Conventional statistical model (logistic regression)
Farrokhi <i>et al.</i> , 2020	Supervised ML algorithms (multiple compared, including gradient boosting machines); trained on 70%, validated on 30%	501 patients; multicenter DBS surgery registry	Any complication; complication within 12 months; return to OR; infection	AUC: any complication 0.86; within 12 months 0.91; return to OR 0.88; infection 0.97	Logistic regression (used to evaluate risk factors)
Chato & Latifi, 2017	Multiple ML methods (SVM, KNN, linear discriminant, tree, ensemble, logistic regression) plus deep learning features extracted by pre-trained CNN	BraTS 2017 dataset; 163 glioma patients with MRI and survival labels (short-, mid-, long-term)	Overall survival classification (three classes: short-, mid-, long-term survivors)	Best accuracy achieved using deep learning features from pre-trained CNN combined with ML classifier (exact figure not reported in abstract)	Conventional 1 ML classifiers without deep learning features
Ive <i>et al.</i> , 2025	NLP (CogStack-MedCAT) + Random Forest classifier trained on extracted clinical concepts from EHR free text	2,268 elective neurosurgery patients; University College London Hospital	Unplanned ITU (intensive therapy unit) admissions post-surgery	Random Forest recall 0.87 (CI 0.82–0.91) for ITU admissions; AUC 0.99; reduced unplanned ITU cases missed by clinicians from 36% to 4%	Human expert clinical judgment (missed 36% of unplanned ITU cases)
Badjatia <i>et al.</i> , 2025	Narrative review of ML approaches (clinical features, neuroimaging, autonomic nervous system metrics)	Review article — no primary dataset	Review of ML for TBI outcome prognostication	Review conclusion: ML integrating multimodal data enhances early detection of deterioration and improves outcome prediction; no single performance figure reported	Traditional prognostication methods

AUC = Area Under the Receiver Operating Characteristic Curve; BraTS = Brain Tumor Segmentation Challenge; CNN = Convolutional Neural Network; DBS = Deep Brain Stimulation; EHR = Electronic Health Record; ITU = Intensive Therapy Unit; KNN = K-Nearest Neighbors; LR = Logistic Regression; MRI = Magnetic Resonance Imaging; NLP = Natural Language Processing; RF = Random Forest; SVM = Support Vector Machine; TBI = Traumatic Brain Injury.

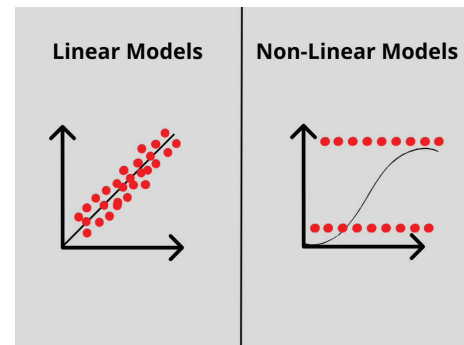


Figure 3: Machine Learning Model Classification Diagram. Static versus dynamic models also contrast logistic regression to ML. Static assumes linear impact between variables, meaning that for the variables' minimum and maximum range, risk can be over- or under-estimated by additive strengths or weaknesses. Nonlinear assumes that demographic, comorbidity, and surgical variables impact one another in a nonlinear approach; at times, certain factors matter less or more depending on other conflicting/intersecting issues, thus distinguishing risk contribution for better accounting of how risk factors impact real-world risk. Created by Raoufian using Canva.

Specific studies in neurosurgery lend credence to this. An extreme gradient boosting model was applied as a precise algorithm to predict complications for intracranial tumor operation patients. Ultimately, from 668 patient data, the model reached a 70% accuracy and an AUC of 0.74, better than conventional statistical models for predictive power.¹⁹ ML has also predicted complications of deep-brain stimulation. Their models were supervised and had high discriminatory performance, with AUC from 0.88 for any complication to 0.97 for infection. These better predictions mean greater flagging of higher risk populations where ML models can more accurately predict preventative measures than conventional means, guiding better preoperative considerations or treatment options.²⁰

These studies show cumulative efforts where MLs and deep learning can analyze the high-dimensional clinical characteristics that are patterned beyond the efforts of standard statistical means, all leading to better postoperative complication prediction and risk of population awareness in neurosurgery for better surgical strategy, patient education, and individualized care.

■ Discussion

Overall, the studies reviewed show that machine learning outperforms current statistical tools in the neurosurgical field. The ability of machine learning to model the interaction of many different variables at once is beyond the capabilities of existing statistical models. Studies by Senders *et al.* have shown that machine learning models outperform statistical models in 30 different studies and in a variety of neurosurgical conditions. Furthermore, the two primary studies reviewed by van Niftrik *et al.* and Farrokhi *et al.* show that these results can also be obtained in tumor and neurosurgical patient populations.

Within these studies, machine learning models based on gradient boosting are among the performing models. Van Niftrik *et al.* discovered that gradient boosting was the performing model for the prediction of complications from intracranial tumors. Farrokhi *et al.* found that gradient boosting was among the performing models for Deep Brain Stimulation procedures.

Ive *et al.* discovered that by utilizing natural language processing to analyze the free-text clinical notes for patients, it was possible to create a model that contained additional information about the patients beyond the structured variables. This new model was able to reduce the percentage of unplanned ICU admissions that were missed by the clinical staff from 36% to 4%. Thus, it would be beneficial for machine learning models to incorporate both structured and unstructured data.

Because machine learning models were found to have such high performance within the clinical field, it is possible to use these models to provide individualized outcome predictions for each patient. Furthermore, these models could be integrated into the current neurosurgical population as a tool to help surgeons make informed decisions about each patient's neurosurgical procedure. While these models may assist in the making of surgeons, the surgeons will still hold complete control and consideration of each patient and the information that cannot be encompassed within any of these models.

■ Weaknesses and Challenges

Single-Center Designs and Limited Generalizability:

The studies by van Niftrik *et al.* and Farrokhi *et al.* each used designs based on data from single institutions. For van Niftrik *et al.*, the model was tested on patients from a single center, with 668 patients included in that test. However, the authors state that external validation would be needed to determine how the model would perform on patients from other centers. Until such validation is performed, the generalizability of the models is not established.

Interpretability:

One of the main drawbacks of the types of models that performed best in both of these studies (gradient boosting methods and deep neural networks) is that they are often difficult to interpret. For instance, it is difficult for a surgeon to understand the reasons that the model provided a risk score to any given patient with a gradient boosting model. Although methods like SHAP and LIME can provide some insight into the decisions of these models, such approaches have yet to be utilized in studies.

Data Quality and Bias:

Because machine learning models learn from the data that is used to train them, the quality of that data can impact the performance of the models. If the training data does not reflect the diversity of the populations that are to be modeled, such as patients of lower socioeconomic status, patients from rural areas, or patients from countries other than Western Europe, the models may have poor performance with those patients. For instance, Ive *et al.* drew their data entirely from a hospital in London.

Ethical and Privacy Concerns:

In order to train these models, vast amounts of sensitive health information must be collected from patients. Additionally, patients' consent must be obtained for the collection of their health information. Finally, the lack of explanations for

the models' predictions presents difficulties for patients and physicians to contest the predictions of the models ethically. Finally, if the training data for these models is biased in some way, the models may present only discriminatory recommendations for certain groups of patients.

■ Conclusion

Machine learning consistently outperforms conventional statistical models in predicting important clinical outcomes following brain tumor resection neurosurgery. Studies by Senders *et al.* report that machine learning models achieve an average area under the curve (AUC) of 0.83 and accuracy of 94.5% in predicting postoperative outcomes, representing a 15-percentage-point advantage over conventional regression models. Subsequent studies have provided additional confirmations of the performance of machine learning models. For instance, Van Niftrik *et al.* found that gradient boosting models achieved an AUC of 0.73 and an accuracy of 0.70 in predicting the development of intracranial tumor complications following resection procedures. Similarly, Farrokhi *et al.* reported AUC values of up to 0.97 in deep brain stimulation machine learning models for the prediction of complications like infection. Ive *et al.* developed natural language processing (NLP) models to achieve similar results with missed admissions to the ICU following neurosurgical procedures, reducing the percentage of missed admissions from 36% to 4%.

Despite the demonstrated performance of machine learning models in brain tumor resection procedures, several challenges must be addressed before such models can be widely adopted into clinical practice. One challenge is validating the models in multiple institutions to ensure the generalizability of the results. Secondly, model performance across studies must be standardized to permit comparison. Thirdly, natural language processing and machine learning models must be explained to clinicians to gain their trust. Finally, data representing diverse populations must be collected to ensure equitable implementation of these models.

When developed to the standards described, machine learning models have the potential to improve the management of patients undergoing brain tumor resection procedures.

■ Acknowledgments

I acknowledge Dr. Sava for her informative information regarding the machine learning process and, of course, Prof. Virgel and Prof. Kay for guiding me through this research paper.

■ References

1. IBM. *What Is Machine Learning?* IBM Think. <https://www.ibm.com/think/topics/machine-learning>
2. Senders, J. T.; Staples, P. C.; Karhade, A. V.; Zaki, M. M.; Gormley, W. B.; Broekman, M. L. D.; Smith, T. R.; Arnaut, O. Machine Learning and Neurosurgical Outcome Prediction: A Systematic Review. *World Neurosurgery* 2018, 109, 476–486.e1. <https://doi.org/10.1016/j.wneu.2017.09.149>. PubMed PMID: 28986230
3. Ive, J.; Olukoya, O.; Funnell, J. P.; Booker, J.; Lam, S. H. M.; Reddy, U.; Noor, K.; Dobson, R. J. B.; Luoma, A. M. V.; Marcus, H. J.

- AI Assisted Prediction of Unplanned Intensive Care Admissions Using Natural Language Processing in Elective Neurosurgery. *npj Digital Medicine* 2025, 8, 549. <https://doi.org/10.1038/s41746-025-01952-0>. PubMed PMID: 40858789
4. Witsch, J.; Siegerink, B.; Nolte, C. H.; Sprügel, M.; Steiner, T.; Endres, M.; Huttner, H. B. Prognostication after Intracerebral Hemorrhage: A Review. *Neurological Research and Practice* 2021, 3 (1). <https://doi.org/10.1186/s42466-021-00120-5>
 5. Jain, S.; Iverson, L. M. Glasgow Coma Scale. *StatPearls*, National Library of Medicine. <https://www.ncbi.nlm.nih.gov/books/NBK513298/>
 6. Campos, A. Hunt and Hess Grading System (Subarachnoid Haemorrhage). *Radiopaedia*. <https://radiopaedia.org/articles/hunt-and-hess-grading-system-subarachnoid-haemorrhage>
 7. Schag, C. C.; Heinrich, R. L.; Ganz, P. A. Karnofsky Performance Status Revisited: Reliability, Validity, and Guidelines. *Journal of Clinical Oncology* 1984, 2 (3), 187–193. <https://doi.org/10.1200/jco.1984.2.3.187>
 8. Perel, P.; Edwards, P.; Wentz, R.; Roberts, I. Systematic Review of Prognostic Models in Traumatic Brain Injury. *BMC Medical Informatics and Decision Making* 2006, 6 (1). <https://doi.org/10.1186/1472-6947-6-38>
 9. Martucci, M.; Russo, R.; Schimperna, F.; *et al.* Magnetic Resonance Imaging of Primary Adult Brain Tumors: State of the Art and Future Perspectives. *Biomedicines* 2023, 11 (2), 364. <https://doi.org/10.3390/biomedicines11020364>
 10. Waite, S.; Grigorian, A.; Alexander, R. G.; *et al.* Analysis of Perceptual Expertise in Radiology—Current Knowledge and a New Perspective. *Frontiers in Human Neuroscience* 2019, 13, 213. <https://doi.org/10.3389/fnhum.2019.00213>
 11. Kamal, V. K.; Agrawal, D.; Kumar, A.; *et al.* Development, Validation and Clinical Utility of Prognostic Models for Patients with Traumatic Brain Injury in an Indian Setting Using Machine Learning and Traditional Approaches: A Study Protocol. *BMJ Open* 2025, 15 (10), e096275. <https://doi.org/10.1136/bmjopen-2024-096275>
 12. Badjatia, N.; Podell, J.; Felix, R. B.; *et al.* Machine Learning Approaches to Prognostication in Traumatic Brain Injury. *Current Neurology and Neuroscience Reports* 2025, 25 (1), 19. <https://doi.org/10.1007/s11910-025-01405-x>. PubMed PMID: 39969697
 13. Breiman, L. Random Forests. *Machine Learning* 2001, 45, 5–32. <https://doi.org/10.1023/A:1010933404324>
 14. Cortes, C.; Vapnik, V. Support-Vector Networks. *Machine Learning* 1995, 20 (3), 273–297. <https://doi.org/10.1007/BF00994018>
 15. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 2016. <https://doi.org/10.1145/2939672.2939785>
 16. Chato, L.; Latifi, S. Machine Learning and Deep Learning Techniques to Predict Overall Survival of Brain Tumor Patients Using MRI Images. *Proceedings of the 2017 IEEE 17th International Conference on Bioinformatics and Bioengineering (BIBE)*, pp. 9–14. <https://doi.org/10.1109/BIBE.2017.00-86>
 17. Park, Y. W.; Yoo, R.-E.; Shin, I.; *et al.* A Multinational Study of Deep Learning-Based Image Enhancement for Multiparametric Glioma MRI. *Scientific Reports* 2025, 15 (1), 32857. <https://doi.org/10.1038/s41598-025-17993-0>. PubMed PMID: 40998920
 18. Hassan, A. M.; Rajesh, A.; Asaad, M.; *et al.* Artificial Intelligence and Machine Learning in Prediction of Surgical Complications: Current State, Applications, and Implications. *The American Surgeon* 2023, 89 (1), 25–30. <https://doi.org/10.1177/00031348221101488>
 19. van Niftrik, C. H. B.; van der Wouden, F.; Staartjes, V. E.; *et al.* Machine Learning Algorithm Identifies Patients at High Risk for Early Complications after Intracranial Tumor Surgery: Registry-Based Cohort Study. *Neurosurgery* 2019, 85 (4), E756–E764. <https://doi.org/10.1093/neuros/nyz145>. PubMed PMID: 31149726
 20. Farrokhi, F.; Buchlak, Q. D.; Sikora, M.; *et al.* Investigating Risk Factors and Predicting Complications in Deep Brain Stimulation Surgery with Machine Learning Algorithms. *World Neurosurgery* 2020, 134, e325–e338. <https://doi.org/10.1016/j.wneu.2019.10.063>. PubMed PMID: 31634625
 21. Bonney, P. A.; Boettcher, L. B.; Burks, J. D.; *et al.* Review of Imaging Modalities for Monitoring Tumor Recurrence after Resection of Malignant Gliomas. *World Neurosurgery* 2017, 98, 686–696. <https://doi.org/10.1016/j.wneu.2016.12.022>
 22. Buchlak, Q. D.; Esmaili, N.; Leveque, J.-C.; *et al.* Machine Learning Applications to Clinical Decision Support in Neurosurgery: An Artificial Intelligence Augmented Systematic Review. *Neurosurgical Review* 2020, 43 (5), 1235–1253. <https://doi.org/10.1007/s10143-019-01163-8>
 23. Wireko, A. A.; Adebuse, F. T.; Wellington, J.; *et al.* Recent Outcomes and Challenges of Artificial Intelligence, Machine Learning and Deep Learning Applications in Neurosurgery. *World Neurosurgery: X* 2024, 100301. <https://doi.org/10.1016/j.wnsx.2024.100301>

■ Author

Arman Raoufian is a freshman at Northwest High School in Maryland with a love for neurosurgery and machine learning. He is particularly focused on developing a deeper understanding of how computational methods can support clinical decision-making and improve patient outcomes. He hopes to major in a medical field.

Quantitative Analysis of Per- and Polyfluoroalkyl Substances (PFAS) in Selective South Bay Water Bodies

Aryanka V. Rekhi

Pioneer High School, 1290 Blossom Hill Road, San Jose, CA 95118, USA; aryankarekhi08@gmail.com

Mentor: Haiqiang Yu

ABSTRACT: Per- and polyfluoroalkyl substances (PFAS) are persistent environmental pollutants, often termed "forever chemicals," known to cause adverse environmental and human health effects. In the San Francisco Bay Area, local surface water bodies and groundwater aquifers serve as critical water supplies, making the monitoring of PFAS concentrations essential for public health protection. While most existing water quality data have predominantly focused on groundwater and aquifers, a detailed quantitative analysis of surface water reservoirs is equally important. Proactive reservoir monitoring provides early detection of contamination before pollutants infiltrate groundwater, where cleanup and remediation are significantly more complex and costly. This study employs targeted quantitative liquid chromatography–tandem mass spectroscopy (LC-MS/MS) of surface waters to analyze a chemical signature dominated by short-chain PFAS compounds, with distributions varying significantly across different water sources. Through advanced analytical chemistry techniques, we present a quantitative assessment of PFAS and related contaminants in selected surface water bodies within the South Bay Area. Furthermore, we correlate these findings with the acceptable concentration limits proposed by environmental regulatory agencies. Ultimately, this research provides actionable, quantitative data to assist local authorities in mitigating PFAS-related contamination in the South Bay Area.

KEYWORDS: Chemistry, Liquid Chromatography, Mass Spectroscopy, Experimental Quantification of Trace PFAS and Related Contaminants, Water Bodies around the San Francisco Bay Area.

■ Introduction

Per- and polyfluoroalkyl substances (PFAS) are a group of persistent, and forever chemicals, and these compounds have a wide range of commercial product applications including industrial polymers, stain repellents, surfactants, waterproofing products, packaging, and aqueous film-forming foams used for firefighting. They are known as "forever chemicals" because they do not break down in the environment.¹ PFAS are often found in river/lake water near industrial manufacturing sites, due to their excellent water solubility and chemical stability posing risks to the ecosystem. They can also accumulate in the human body over time, leading to adverse human health effects, especially related to the liver and immune system.

The Bay Area/Silicon Valley is the center of technology and innovation, particularly in software, hardware, and semiconductors. Additionally, it is also a major hub for various other industries such as medical device manufacturing, biotechnology, and life sciences. Industrial discharge has been reported to contribute to the release of PFAS compounds, which could contaminate the nearby water bodies.² The South Bay is geographically different from the Central and North Bay Area and has limited tidal flushing. This can increase the residence time of contaminants, and SFEI shows elevated PFAS concentrations linked to the region because of manufacturing, industrial, municipal, and fire training sites. Because a major portion of the local water supply comes from surface water bodies and groundwater, it is very critical to understand and determine the level of PFAS-related compounds in these water bodies and understand the overall extent of PFAS exposure in the eco-

system. We conducted a quantitative analysis of various South Bay reservoirs to reveal elevated total PFAS concentrations and the distribution of these contaminants, aiming to study a direct correlation with their proximity to urban and industrial point sources.

This study is significant because much of the available water quality data in the Bay Area has historically focused on groundwater and aquifer contamination.³ The long-chain PFAS compounds (containing eight or more carbon atoms) are highly hydrophobic (water-repelling), causing them to partition out of the water system and absorb into the sediments. Conversely, the short-chain compounds are highly water-soluble and are highly mobile in aquatic environments.

Although the surface water in these reservoirs may not be directly suitable for drinking, it is more vulnerable to PFAS contamination than well water.⁴ Furthermore, these reservoirs remain interconnected components of the broader water supply system; monitoring them is critical because they serve as early indicators of contaminants before they migrate into the groundwater supply, where remediation efforts become significantly difficult and costly. Therefore, we hypothesized that a targeted quantitative LC-MS/MS analysis of surface water would reveal a chemical signature dominated by short-chain PFAS compounds, and their distribution would vary significantly across different water sources.

In this study, we collected water samples from several rivers/lakes around the SF Bay Area and used LC-MS/MS to identify 12 different types of PFAS compounds. These 12 PFAS compounds are well studied, widely used in the industry, and

heavily regulated by the EPA, the Environmental Protection Agency,⁵ and the California Water Resources Control Board.⁶ Currently, the EPA does not have a single enforceable standard for all rivers and lakes, but has established enforceable drinking water standards, called MCLs Maximum Contaminant Levels,⁷ for 5 specific PFAS chemicals. For drinking water, PFOA and PFOS have an MCL of 4.0 ppt (parts per trillion), while PFNA and PFHxS have an MCL of 10.0 ppt, and for PFHxA, the EPA does not currently have a finalized federal MCL. For protecting aquatic life, the EPA has established non-enforceable benchmark values for PFOA and PFOS, which are 4.0 ppt for PFOA and 8.0 ppt for PFOS.

LC-MS/MS is widely used for PFAS analysis due to its high sensitivity and selectivity at trace concentration levels.^{8,9} Given the low detection limits required for PFAS in water, precautions were taken to reduce system-derived contamination, including the use of a trap column and large-volume injections to ensure reliable quantification. Due to the low detection limit requirements (parts per trillion) of PFAS in water, there have been a lot of analytical challenges due to potential contamination from lab equipment and reagents. In this study, we used a trap column after the LC pump to prevent potential contamination from the system and large-volume injection to achieve the targeted sensitivity needed for the PFAS quantification.

■ Methods

Thousands of different PFAS exist, are used in countless products, and are now present in the environment. Table 1 lists the most prominent PFAS compounds along with their associated information.

Table 1: Detailed list of PFAS compounds. This table summarizes the 14 per- and polyfluoroalkyl compounds to be analyzed, spanning C4 to C14 chain lengths with molecular weights from 214.03 to 714.11 g/mol.

Compound Code	Full Name	MW g/mole	Formula	CAS
PFPeA	Perfluoropentanoic acid	263.9	C ₅ HF ₉ O ₂	2706-90-3
PFHxA	Perfluorohexanoic acid	314.1	C ₆ HF ₁₁ O ₂	307-24-4
PFHpA	Perfluoroheptanoic acid	364.1	C ₇ HF ₁₃ O ₂	375-85-9
PFNA	Perfluorononanoic acid	464.1	C ₉ HF ₁₇ O ₂	375-95-1
PFDA	Perfluorodecanoic acid	514.1	C ₁₀ HF ₁₉ O ₂	335-76-2
PFUnDA	Perfluoroundecanoic acid	564.1	C ₁₁ HF ₂₁ O ₂	2058-94-8
PFDoDA	Perfluorododecanoic acid	614.1	C ₁₂ HF ₂₃ O ₂	307-55-1
PFBS	Perfluorobutane sulfonic acid	300.1	C ₄ HF ₉ O ₃ S	375-73-5
PFOS	Perfluorooctane sulfonic acid	500.1	C ₈ HF ₁₇ O ₃ S	1763-23-1
PFOA	Perfluorooctanoic acid	414.1	C ₈ HF ₁₅ O ₂	335-67-1
PFTA	Perfluorotetradecanoic acid	714.1	C ₁₄ HF ₂₇ O ₂	376-06-7
PFTDA	Perfluorotridecanoic acid	664.1	F ₃ C(CF ₂) ₁₁ CO ₂ H	72629-94-8
PFHxS	Perfluorohexane sulfonic acid potassium salt	438.2	C ₆ F ₁₃ KO ₃ S	3871-99-6
PFBA	Perfluorobutanoic acid	214.0	C ₄ HF ₇ O ₂	375-22-4

Methanol (LC-MS/MS Grade, Fisher Scientific, Lot number: 222630), Water (LC-MS/MS grade, Fisher Scientific, Lot number: 242945), Trap Column (Thermo Scientific BDS Hypersil Gold (50x2.1 mm, 5 μ, Lot number: 21524), Analytical Column (Thermo Scientific Hypersil Gold 100x2.1 mm, 1.9 μ, Lot number: 21565) and LC vials (Catalog number: 03452223, Lot number: TSKM4D176GF). All the standards

with a specific concentration were diluted with LC-MS/MS-grade water.

Our study focuses on the water samples collected from different water bodies in the SF Bay Area. Table 2 lists the names of lakes/ rivers and the sources of water samples. Most of these water reservoirs are man-made, or significant portions of their watershed are altered by man-made features. Based on our understanding, Vasona Lake/ Reservoir is managed by Valley Water, and it is a source that feeds the drinking water system after treatment. Almaden Reservoir is one of the Santa Clara Valley Water District's 10 reservoirs and supplies water to drinking water plants from groundwater recharge. Calero Reservoir captures runoff and provides water directly to drinking water plants. The others are recreation, flood control, or groundwater recharge.

Table 2: Name and location of reservoirs (water sample sources) chosen for this study. Nine water bodies were selected across the SF Bay Area to capture a gradient of urban, recreational, and drinking-water supply reservoirs.

Sample Source	Location
Vasona Lake (57 acres)	Los Gatos, CA
Lake Cunningham (50 acres)	East San Jose, CA
Lake Elizabeth (83 acres)	Fremont, CA
Guadalupe Creek (10 miles)	Almaden San Jose, CA
Cottonwood Lake (~13 acres)	San Jose, CA
Sandy Wool Lake (~14 acres)	Milpitas, CA
Almaden Reservoir (62 acres)	South San Jose, CA
Calero Reservoir (349 acres)	Almaden valley CA
Lake Almaden (32 acres)	South San Jose, CA

According to existing literature, various inorganic pollutants, including heavy metals and Biotoxins, may exist in these reservoirs; however, our study is solely focused on the PFAS compounds. We collected the water samples in properly labeled, clean glass bottles; all the water samples were centrifuged, and the supernatants were transferred to a 1.5 mL LC autosampler vial.

A Vanquish UHPLC system equipped with a special PFAS Kit, including a solvent filter and solvent tubing, and coupled with a TSQ Altis Plus (Thermo Fisher Scientific, San Jose, California) and a heated electrospray ion source (HESI), was used to acquire SRM (Selected Reaction Monitoring) mass spectra in negative mode. The HESI source parameters were set with an ion spray voltage of 2.8 kV (negative ionization mode); sheath gas flow rate of 40 arbitrary units, and auxiliary gas flow rate of 10 arbitrary units; the temperatures of the vaporizer and capillary were both kept at 325 °C. Table 3 lists the detailed SRM methods for targeted compounds.

Table 3: Detailed SRM method parameters for targeted compounds. Chromatographic separation was performed on a Thermo Hypersil Gold C18 column with a BDS Hypersil Gold trap column at 0.25 mL/min using water (mobile phase A) and methanol with 10 mM ammonium acetate and 0.1% acetic acid (mobile phase B). The gradient was 0-30% B (0-1 min), 30-45% B (1-6 min), 45-80% B (6-13 min), 80-95% B (13-14 min), and 95% B hold (14-18 min), followed by a 7 min equilibrium. The injection volume was 10 μL.

Compound	Retention Time (min)	RT Window (min)	Precursor (m/z)	Product (m/z)	Collision Energy (V)	RF Lens (V)	Min Dwell Time (ms)
PFBA	4.84	4	212.98	168.97	9	30	73.05
PFPeA	7.58	4	262.98	219.04	9	31	48.38

PFBS	8.26	4	298.94	79.96	34	116	28.52
	8.26	4	298.94	98.95	29	116	28.52
PFHxA	10.17	4	312.97	63.19	15.46	39	15.28
	10.17	4	312.97	119.04	18.76	39	15.28
	10.17	4	312.97	268.97	9	39	15.2
PFHpA	11.79	4	362.97	119.05	19.52	43	11.12
	11.79	4	362.97	168.97	15.53	43	11.12
	11.79	4	362.97	319.04	9	43	11.12
PFHxS	11.92	4	398.93	79.95	39	135	11.12
	11.92	4	398.93	98.95	35	135	11.12
PFOA	12.93	4	412.96	169	16.1	49	10.63
	12.93	4	412.96	219	14.55	49	10.63
	12.93	4	412.96	369.04	9	49	10.63
PFNA	13.82	4	462.96	169	17.51	52	10.63
	13.82	4	462.96	219.01	15.23	52	10.63
PFOS	13.82	4	462.96	418.97	9	52	10.63
	13.85	4	498.93	79.96	47	159	10.63
	13.85	4	498.93	98.95	40	159	10.63
	13.85	4	498.93	129.95	20	159	10.63
PFDA	14.56	4	512.96	219.01	16.14	56	10.63
	14.56	4	512.96	269.04	15.8	56	10.63
	14.56	4	512.96	469.04	9	56	10.63
PFUnDA	15.16	4	562.95	219	17.32	62	10.63
	15.16	4	562.95	269.03	16.94	62	10.63
	15.16	4	562.95	518.97	9	62	10.63
PFDoDA	15.62	4	612.95	169.03	23.69	67	10.63
	15.62	4	612.95	319.04	17.54	67	10.63
	15.62	4	612.95	569	9	67	10.63
PFTrDA	15.9	4	662.95	168.97	25.16	71	10.63
	15.9	4	662.95	369.07	17.85	71	10.63
	15.9	4	662.95	619.04	9	71	10.63
PFTeDA	16.1	4	712.94	319.05	19.86	74	10.63
	16.1	4	712.94	369.04	18.87	74	10.63
	16.1	4	712.94	668.97	9	74	10.63

All data acquisition and analysis were conducted by Xcalibur 4.5 and TraceFinder 5.1. Samples were separated on a Thermo Hypersil Gold C18 column, and the mobile phase consisted of water (mobile phase A) and 10 mM ammonium acetate, 0.1% acetic acid in MeOH (mobile phase B: 0.078 g ammonium acetate and 1 mL acetic acid were added into 1000 mL methanol). The gradient is from 0% B to 30% B in 0–1 min, from 30% B to 45% B in 1–6 min, from 45% B to 80% B in 6–13 min, from 80% to 95% in 13–14 min and hold at 95% B for 4 min for a total of 18 min run with the flow rate of 0.25 mL/min. The system was equilibrated with 95% A for 7 min before the next injection. The sample injection volume for water samples and standard samples from the lakes is 10 µL. A trap column was used right after the LC pump and before the autosampler to prevent system contamination and increase sensitivity.

■ Results and Discussion

To evaluate chromatographic separation and detection of PFOA, PFOS, PFNA, PFHxS, and PFHxA, and the suitability of the method, the chromatograms of five PFAS standards at 4 ng/L made in LC-MS/MS grade water were measured and are shown in Figure 1.

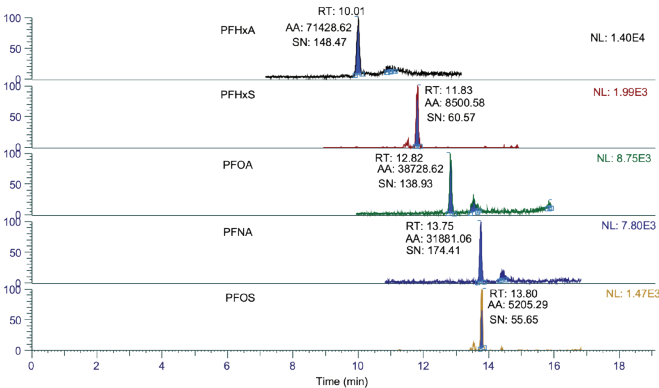


Figure 1: Representative LC-MS/MS chromatogram of PFHxA, PFHxS, PFOA, PFNA, and PFOS spiked at 4 ng/L in LC-MS/MS grade water. Peak retention time (RT), integrated peak area (AA), and signal-to-noise (S/N) were used to assess chromatographic performance and detection sensitivity. All five target compounds show baseline chromatographic separation with retention times between 10 and 15 minutes. Signal-to-noise (S/N) ratios for all analytes exceeded 55 at 4 ng/L, demonstrating method sensitivity well above the typical S/N = 10 quantification threshold and suitability for low-level environmental monitoring.

The peak identification, along with retention times, peak area, and signal-to-noise ratios information derived from the chromatogram in Figure 1, is listed in Table 4. All the analytes are well-separated and detected within the retention time window between 10 min and 15 min, and the peak signal/noise ratios are 55.65-148.47. The method achieved the necessary sensitivity to meet the reportable threshold for all 5 analytes. All the collected water samples were directly injected into the LC-MS/MS system with the targeted SRM method to conduct experimental quantitative screening of PFOA, PFOS, PFNA, PFHxS, and PFHxA across all water samples from different lakes and rivers. The raw data were processed and analyzed with TraceFinder software. We have used the peak area of a particular PFAS compound and compared it with the same compound with a known concentration in ultra-pure PFAS-free water. Before the water sample analysis, a mixture of PFAS standards was run with the same method, and the retention time data were summarized in Table 4. The screening of 14 PFAS compounds in water samples was conducted using the retention time of the individual PFAS compounds listed in Table 3. A peak from the raw data at a similar retention time to the standard mixture data indicated a positive hit. No peak from the raw data indicated a negative hit. The concentration determination of the positive hit was performed using direct comparison of the peak areas of unknown samples and standard samples.

Table 4: Peak Identification and peak parameters of PFHxA, PFHxS, PFOA, PFNA, and PFOS compounds. Target analytes were identified by comparing retention times with those of the standard reference mixture. Quantification was performed by direct comparison of peak areas between samples and corresponding external standards.

Peak No.	Compound Name	Retention Time (min)	Peak Area	Signal/Ratio
1	PFHxA	10.01	71428	148.47
2	PFHxS	11.83	8500	60.57
3	PFOA	12.82	38728	138
4	PFNA	13.75	31881	174
5	PFOS	13.80	5205	55.65

The example chromatogram of water samples from Lake Cunningham is depicted in Figure 2. The chromatographic analysis indicated positive hits for PFHxA, PFHxS, and PFOA and negative hits for PFNA and PFOS, with no peaks at the expected retention time. The concentrations of PFHxA, PFHxS, and PFOA were determined based on the peak area. A detailed summary of PFAS occurrence across all nine reservoirs is listed in Table 5.

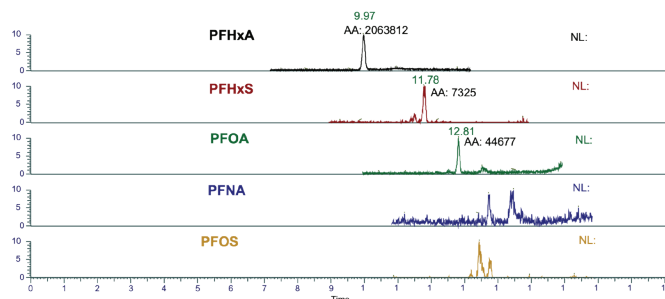


Figure 2: Chromatogram of water sample from Lake Cunningham. The chromatographic analysis revealed the presence of PFHxA, PFHxS, and PFOA, which were positively identified, whereas there were negative hits for PFNA and PFOS with no peaks at expected retention times. The concentrations of the detected analytes were determined based on integrated peak areas.

Table 5: Estimated concentration of different PFAS compounds in collected water samples measured across different collection trials within the same lake. Chromatographic analysis positively identified PFHxA, PFHxS, and PFOA, while PFNA and PFOS were not detected. PFHxA dominated the profile, indicating short-chain PFAS as major contributors at this site. Concentrations of identified analytes were quantified based on integrated peak areas. The estimated concentrations were determined based on the comparison between integrated peak areas from each water sample and the peak areas from the standards at 4.0, 10.0, 25.0, 50.0, and 100.0 ng/L. The method detection limit for all 5 standards is 4 ng/L.

	PFHxA (ng/L)	PFHxS (ng/L)	PFOA (ng/L)	PFNA (ng/L)	PFOS (ng/L)
Vasona Lake	4-10	< 4	< 4	< 4	< 4
Lake Cunningham	25-50	< 4	4-10	< 4	< 4
Lake Elizabeth	25-50	< 4	< 4	< 4	< 4
Guadalupe Creek	25-50	< 4	< 4	< 4	< 4
Cottonwood Lake	25-50	< 4	< 4	< 4	< 4
Sandy Wool Lake	50-100	< 4	< 4	< 4	< 4
Almaden Reservoir	25-50	< 4	< 4	< 4	< 4
Calero Reservoir	25-50	< 4	< 4	< 4	< 4
Lake Almaden	25-50	< 4	< 4	< 4	< 4

Detection Limit: 4 ng/L

(All values reported by values < 4 indicate concentrations below the method reporting limit of 4 ng/L)

Based on several city reports and government resources, Lake Cunningham is facing more ecological and water quality issues than merely PFAS contamination. However, our analysis indicates a higher level of PFOA in Lake Cunningham that may be linked to its location in a highly developed urban region in the surrounding hills, where stormwater runoff percolated through the underground gravel and flowed into Thompson Creek and Lake Cunningham. While it could be challenging to identify the exact source of PFOA, it is widely recognized that the common sources of PFOA within the water system may include firefighting foams, industrial waste, biosolids, and various consumer product industries. Poor circula-

tion and flooding issues allow the contaminants to be trapped in the sediments for a long period of time.

In terms of regulatory context, PFOA is categorized by the International Agency for Research on Cancer (IARC) as Group 1, “carcinogenic to humans”. This classification is based on sufficient evidence of carcinogenicity in animal studies and strong mechanistic evidence in humans, although there is limited evidence for renal cell carcinoma and testicular cancer.¹⁰ Additionally, some epidemiological studies have reported associations with other cancer sites, such as thyroid, breast, and ovarian cancer; however, IARC has concluded that the current evidence on humans from these cancer sites is insufficient.

Although the overall levels of PFAS material may fall within acceptable limits, localized areas of elevated concentration can still be present. A higher concentration of pollutants can significantly affect groundwater, as groundwater and surface water are inseparably linked components of the water cycle and facilitate the exchange of chemicals. The water on the surface could seep downwards, thereby replenishing the underlying aquifer, where the groundwater can stay longer for extended periods. Over time, this groundwater can cycle back into surface bodies such as streams and wells. The duration and concentration for which the water stays underground, along with the chemicals, can significantly alter the quality of the water. Consequently, the surface water has the potential to degrade the overall water quality of the ecosystem.

The presence of harmful substances in the environment underscores the critical need for effective management strategies to protect both the ecosystem and public health. In response, the authorities have implemented several measures, such as regulating stormwater and runoff, pollution prevention practices, reducing plastic materials, especially BPA-free, community training and involvement, and student education, such as Bio Site and Environmental Corps programs. The South Bay Water Department is also starting an ion exchange technology plant to eliminate PFAS and treat groundwater.¹¹ Alternative technologies, such as reverse osmosis and filtration, can effectively remove long-chain contaminants; however, increased awareness is needed to ensure the safety of surface water by restricting the source of contaminants into reservoirs and safeguarding waterways.

■ Conclusion

The results support the study's hypothesis that surface waters exhibit a chemical signature dominated by highly mobile, short-chain PFAS compounds (specifically PFHxA) rather than long-chain variants like PFOS. The elevated levels of PFOA in Lake Cunningham may be associated with urban stormwater runoff. While the overall levels of PFAS detected generally fall within acceptable regulatory limits, the study emphasizes the critical importance of continuous surface water monitoring. Because surface waters are interconnected with groundwater aquifers, these reservoirs serve as an essential early warning system for contamination before pollutants migrate into groundwater supplies, where remediation becomes substantially more complex and costly. The spatial analysis showed a clear link, indicating that PFAS levels drop as we move away

from urban locations. Locations within 1 km of busy developments or major highways had total PFAS concentrations that were 3.2 – 5.8 times greater than reference sites in less-developed areas. Recent studies have shown that concentration decreases by about 60% for every tenfold increase in distance from urban sources.¹² This quantitative relationship backs up the idea that stormwater runoff from hard surfaces is a significant way for PFAS to enter the environment and suggests that proximity to urban land use is a strong indicator of local PFAS levels.

These results are crucial for managing watersheds in areas that are becoming more urbanized. Additional studies could be conducted to provide a direct mechanism for stormwater runoff and to investigate the short-distance contamination gradient in high-density urban, suburban, and industrial areas compared to reference watersheds.

■ Acknowledgments

The authors are very grateful to Dr. Laura Bald of Pioneer High School for reviewing this paper and providing valuable feedback.

■ References

1. Al-Maqtari, Q. A.; Mahdi, A. A.; Othman, N.; Noman, A. E.; Alsubhi, L. M.; AlAnsi, W.; Asharuddin, S. M.; Talib, M. K.; Supramaniam, S.; Al-Gheethi, A. Comprehensive review of per- and polyfluoroalkyl substances in drinking water: Physiochemical properties, sources, health impacts, regulatory limits, and future perspectives. *Results Eng.* 2025, 28, 107169. DOI: 10.1016/j.rineng.2025.107169
2. Petre, M.-A.; Salk, K.R.; Stapleton, H.M.; Ferguson, P.L.; Tait, G.; Obenour, D.R.; Knappe, D.R.U., & Genereux, D.P. Per- and polyfluoroalkyl substances (PFAS) in river discharge: Modeling loads upstream and downstream of a PFAS manufacturing plant in the Cape Fear watershed, North Carolina. *Science of The Total Environment*, 2022, 831, 154763. DOI: <https://doi.org/10.1016/j.scitotenv.2022.154763>
3. Dao, P. U.; Heuzard, A.; Le, T. X. H.; Zhao, J.; Yin, R.; Shang, C.; Fan, C. The impacts of climate change on groundwater quality: A review. *Sci. Total Environ.* 2024, 912, 169241. DOI: 10.1016/j.scitotenv.2023.169241
4. Peter, L. G.; Lee, L. S. Sources and Pathways of PFAS occurrence in water sources: Relative contribution of land-applied biosolids in an agriculturally dominated watershed. *Environ. Sci. Technol.* 2025, 59 (2), 1344–1353. DOI: 10.1021/acs.est.4c09490
5. U.S. Environmental Protection Agency. Per- and Polyfluoroalkyl Substances (PFAS), Final PFAS National Primary Drinking Water Regulations. <https://www.epa.gov/sdwa/and-polyfluoroalkyl-substances-pfas> (accessed 2025-12-04).
6. California State Water Resources Control Board, Division of Drinking Water. PFAS: Per- and Polyfluoroalkyl Substances Drinking Water Systems. https://www.waterboards.ca.gov/drinking_water/certlic/drinkingwater/pfas.html (accessed 2025-11-20).
7. Page, G. W.; Greenberg, M. Maximum Contaminant Levels for Toxic Substances in Water: A Statistical Approach. *J. Am. Water Resour. Assoc.* 1982, 18 (6), 955–963. DOI: 10.1111/j.1752-1688.1982.tb00101.x
8. Nahar, K.; Zulkarnain, N. A.; Niven, R. K. A Review of Analytical Methods and Technologies for Monitoring Per- and Polyfluoroal-

kyl Substances (PFAS) in Water. *Water* 2023, 15, 3577. <https://doi.org/10.3390/w15203577>.

9. Coggan, T. L.; Anumol, T.; Drage, D.; Shimko, K.; Stevenson, G.; Roberts, D. T. A Comprehensive Analytical Workflow for the Detection of over 60 Per- and Polyfluoroalkyl Substances (PFAS) in Diverse Environmental and Biological Matrices. *Anal. Bioanal. Chem.* 2019, 411, 6841–6855. <https://doi.org/10.1007/s00216-019-02080-6>.
10. American Cancer Society. PFOA, PFOS, and Related PFAS Chemicals. <https://www.cancer.org/cancer/risk-prevention/chemicals/teflon-and-perfluorooctanoic-acid-pfoa.html> (accessed 2024-05-31).
11. San Jose Water Company. Williams Road Station PFAS treatment facility Project (CP24-025 and ER24-166): initial study/mitigated negative declaration; City of San Jose: San Jose, CA, 2025. <https://www.sanjoseca.gov/Home/Components/News/News/6804/4699> (accessed 2025-07-18).
12. Wang, H.; Wang, Y.; Zhao, H.; Zhang, J.; Li, S. Urban soils encounter more severe per- and polyfluoroalkyl substance contaminations and associated exposure risks. *Environ. Pollut.* 2026, 127653. DOI: 10.1016/j.envpol.2026.127653.

■ Author

Aryanka V. Rekhi is currently a junior in high school located in San Jose, California. Alongside her core academic studies, she has been actively involved in diverse research projects focused on environmental chemistry, computational neuroscience, and the impact of GenAI in the academic environment. Her deep interest in medicine has driven her to explore various research opportunities. She is also pursuing a degree in classical dance, which she expects to complete by the summer of 2026. In her spare time, she enjoys volunteering and traveling to immerse herself in different cultures and experiences.

Evaluation of Nitrate Concentration Estimation using Machine Learning Models in Lake Erie Watersheds, USA

Saeyool Chloe Park

Seoul International School, 15 Seongnam-daero 1518beon-gil, Sujeong-gu, Seongnam-si, Gyeonggi-do, 13113 Republic of Korea;
chloe.park.0405@gmail.com

ABSTRACT: This study analyzes machine-learning methods used to estimate daily nitrate concentrations by analyzing streamflow and water quality data from 18 watersheds in the Lake Erie area. Daily nitrate was predicted by using 10 input variables: Julian date, streamflow, total suspended solids, chloride, sulfate, total phosphorus, soluble reactive phosphorus, total Kjeldahl nitrogen, total silica, and conductivity. There are 27 machine learning models that are compared using K-fold cross-validation and analyzed with the mean absolute error (MAE), root mean square error (RMSE), and coefficient of determination (R^2). But agricultural land use also dominates 15 of the 18 analyzed watersheds, while two watersheds are urban-dominated, and one (Muskingum) watershed shows the highest forest land use. As a result, Rational Quadratic Gaussian Process Regression (RQ-GPR) and Random Forest (RF) among the machine learning models have the strongest overall performance. The correlation analysis of the variables shows that streamflow and Julian date are the most important variables when it comes to predicting nitrate concentration. This research supports the availability of machine learning models for estimating nitrate concentration, the water quality variable.

KEYWORDS: Julian Date, Lake Erie. Machine Learning, Nitrate Concentration, Streamflow.

■ Introduction

Nitrate is one of the major water quality variables because it is used in river impairment, eutrophication, and downstream ecological effects.^{1,2} In agricultural watersheds, nitrate concentrations are significantly affected by hydrologic streamflow, land-use practices, and seasonal biogeochemical controls.³⁻⁶ However, directly and continuously monitoring the nitrate concentration is more difficult and time-consuming than collecting routine hydrologic or supporting water quality measurements. So, the observed nitrate concentration during the monitoring period is usually collected on a weekly or monthly basis. But, the data were not enough to estimate the annual nitrate loads, leading to inconclusive results.⁷ Therefore, it was determined that utilizing machine learning models to estimate daily nitrate concentration offers a faster and simpler approach to daily monitoring.^{8,9} In addition, due to watershed heterogeneity, nonlinear models have been widely studied because nitrate concentrations are affected by land use and watershed characteristics.^{10,11} As a result, it is necessary to apply machine learning models to estimate nitrate concentration because they estimate nonlinear relationships in data very well. This leads to the main objectives of this study: (1) to construct a nitrate concentration estimation method using selected observed nitrate concentrations with hydrologic streamflow, water-quality, and time variables; (2) to compare the estimated nitrate concentration of 27 machine learning models with the observed nitrate, where the concentration evaluates the performance of machine learning models with the experimental watersheds; and lastly (3) to identify the variables with the greatest effect on nitrate estimation.

■ Materials and Methods

Study area and watershed characteristics:

18 watersheds from Lake Erie, in Ohio, USA, were used for this study, including Maumee, Wolf, Tiffin, Lost, West, Rock Creek, Beaver, Cuyahoga, Huron, Muskingum, Portage, Raisin, Scioto, Blanchard, Chickasaw, Potato Run, Shallow Run, and S. Turkeyfoot. Figure 1 shows the location of the applied watersheds, while Figure 2 presents the land-use characteristics of the watersheds used in this study. Agriculture was the dominant land use in 15 of the 18 watersheds. In Figure 2, the Wolf and Cuyahoga watersheds were the only urban-dominated basins, with urban fractions of 43.9% and 39.5%. Muskingum had the greatest forest fraction (43.0%), while West (92.0%) and S. Turkeyfoot (90.3%) had the highest agriculture fractions.

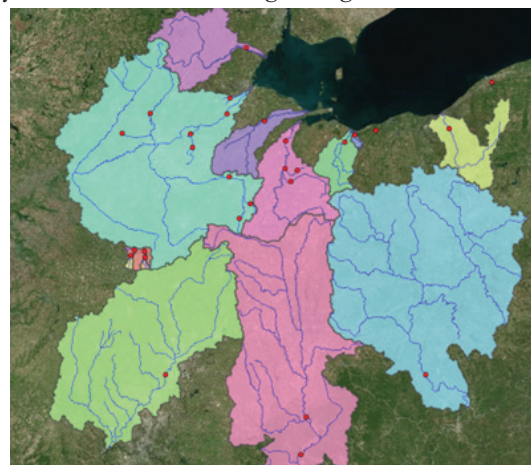


Figure 1: Locations of the 18 watersheds in the Lake Erie region, Ohio, USA, used in this study. Watershed boundaries and monitoring locations are shown for the selected basins.

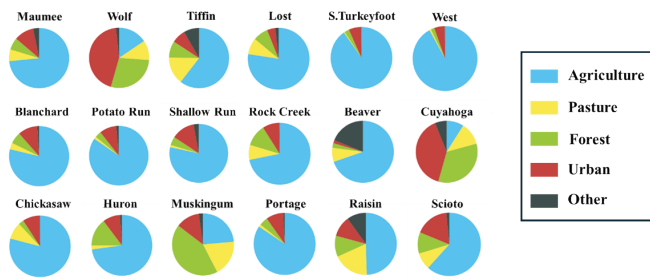


Figure 2: Land-use composition of the 18 watersheds used in this study. Agricultural, pasture, forest, urban, and other land-use categories are presented as percentages of total watershed areas.

Data sources and predictors:

Daily streamflow data were collected from the U.S. Geological Survey (USGS) (<https://waterdata.usgs.gov>), and also the water-quality observations were collected from the National Center for Water Quality Research (NCWQR) (<https://ncwqr-data.org>). To match the time period of the streamflow and water quality data, the monitoring periods in the watershed datasets ranged from 1974-10-02 to 2023-05-15, depending on the watershed. The monitoring records were arranged into a daily database. The variables collected in this study, listed in Table 1, include Julian date (day of year), streamflow, total suspended solids (TSS), chloride, sulfate, total phosphorus (TP), soluble reactive phosphorus (SRP), total Kjeldahl nitrogen (TKN), total silica, and conductivity. In addition, Table 1 also presents the characteristics of the adopted variables. In the end, the target simulation variable to be measured was daily nitrate concentration.

Table 1: Input variables of machine learning models for daily nitrate estimation.

Variables	Unit/form	Characteristics
Julian date	day of year	Seasonality and recurring annual controls
Streamflow	m ³ /s	Hydrologic flow effects
TSS	mg/L	Soil water quality parameter
Chloride	mg/L	Chemical water quality parameter
Sulfate	mg/L	Chemical water quality parameter
TP	mg-P/L	Nutrient water quality parameter
SRP	mg-P/L	Nutrient water quality parameter
TKN	mg-N/L	Nutrient water quality parameter
Total silica	mg-Si/L	Chemical water quality parameter
Conductivity	Umho	Soil water quality parameter

Model comparison framework:

First of all, Figure 3 shows the overall study process. 10 variables were collected as observational data, including time, two soil, three chemical, and three nutrient water qualities, and hydrologic variables, and 27 machine learning models were used to estimate daily nitrate concentration. The 27 machine learning models were then further categorized into nine machine learning groups according to Table 2. The categories/groups are: linear regression models, regression trees, support vector machines (SVMs), Gaussian process regression (GPR), neural networks, kernel approximation methods, ensembles of trees, efficiently trained linear models, and random forests (RFs). Eventually, all 27 models and the categories were used to compare watershed performance. Furthermore, the created dataset was split into 80% for training and 20% for testing, with 10-

fold cross-validation used during model application. Then, the model performance was analyzed with MAE, RMSE, and R2. Finally, the performance of machine learning models was compared, and the optimal models to produce the best results were presented under the general conditions.

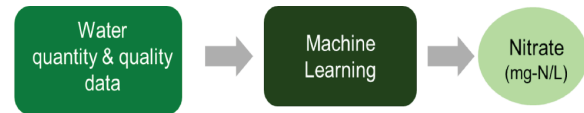


Figure 3: Overall framework of this study, including input variable data (Table 1) collection, machine learning model application (Table 2), and nitrate concentration estimation (Figures 4-7).

Table 2: Machine learning categories and models screened in the MATLAB-based application.

Category	Representative models	Count
Linear regression models	Robust, linear, interaction, and stepwise linear regressions	4
Regression trees	Fine, medium, and coarse trees	3
Gaussian Process Regression (GPR)	Rational quadratic GPR, squared exponential GPR, Matern 5/2 GPR, and exponential GPR	4
Support vector machines (SVMs)	Linear SVM, quadratic SVM, cubic SVM, fine Gaussian SVM, medium Gaussian SVM, and coarse Gaussian SVM	6
Neural networks	Narrow, medium, and wide neural networks	3
Kernel approximation models	SVM kernel and least-squares kernel regressions	2
Ensembles of trees	Boosted trees and bagged trees	2
Efficiently trained linear models	Efficient linear least squares and efficient linear SVMs	2
Random forest (RF)	Random forest	1

Performance evaluation methods:

Three standard regression error analysis methods were utilized: root mean square error (RMSE), mean absolute error (MAE), and coefficient of Determination (R2) to evaluate machine model performance. These methods compare the predicted nitrate concentrations with the observed values.

The RMSE quantifies the standard deviation between the estimated nitrate concentrations by machine learning models and the observed nitrate concentrations as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (1)$$

where y_i is the observed nitrate concentration, $\hat{y}_i$ is the simulated nitrate concentration by machine learning models, and n is the number of data points.

The MAE provides a measure of average nitrate estimation bias as follows:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (2)$$

The R2 quantifies the proportion of variance in observed and simulated nitrate concentrations by machine learning models, as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (3)$$

where $\bar{y}$ is the mean of the observed nitrate concentrations. While the ranges of the RMSE and MAE are data-dependent, ranging from zero to infinity, zero representing the optimal value, the R2 range ranges from 0 to 1, 1 representing the optimal value. RMSE presents the magnitude of the estimation errors. A smaller RMSE value indicates the machine learning models produce fewer large mistakes and create a more

accurate estimation. MAE shows the average absolute difference between the estimated and observed values. Lower MAE values demonstrate a better machine learning model performance because the estimations are closer to the actual data. R2 measures how well the model explains the changes in nitrate concentration. The R2 value ranges from 0 to 1, and values closer to 1 indicate that the machine learning model estimates the observed data more accurately. Overall, a well-performing model provides low RMSE and MAE and a high R2.

Results

Watershed-wise machine learning models evaluation:

Figure 4 shows the performance results of 27 machine learning models that estimate nitrate concentration, using 10 input variables per watershed, and also combining RMSE, MAE, and R2. The Rational Quadratic Gaussian Process Regression (RQ-GPR) had the best-performing model compared to all three evaluation methods in eight watersheds, including Cuyahoga, Potato Run, Muskingum, Scioto, Shallow Run, S. Turkeyfoot, West, and Wolf. Meanwhile, a small number of the bagged trees model and a medium neural network were selected. The RF dominated all three evaluation methods in the Raisin and Maumee watersheds. Lastly, in the Huron, Rock Creek, Portage, and Tiffin watersheds, the Exponential GPR and RQ-GPR models showed the best performance results. Overall, in Figure 5, five of 27 different machine learning models were shown as the optimal models depending on watersheds and performance methods. Among them, RQ-GPR was selected as the most optimal machine learning model in each performance method per watershed. RQ-GPR appeared in 30 out of the 54 optimal performance-error methods, and RF was the second most frequently selected machine learning model, with 14 selections, and, in addition, the exponential GPR with only 6 selections. These results suggest that RQ-GPR is the most powerful model for nitrate concentration, whereas RF performs especially well in some agricultural watersheds.

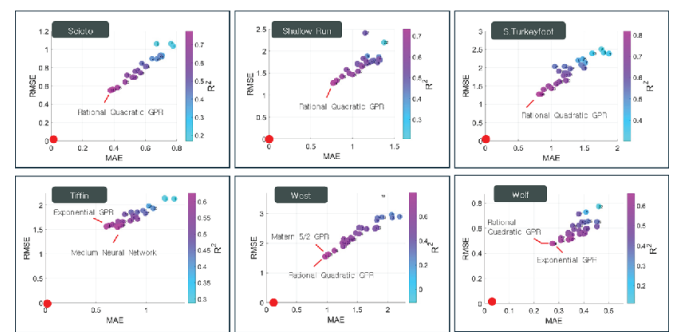
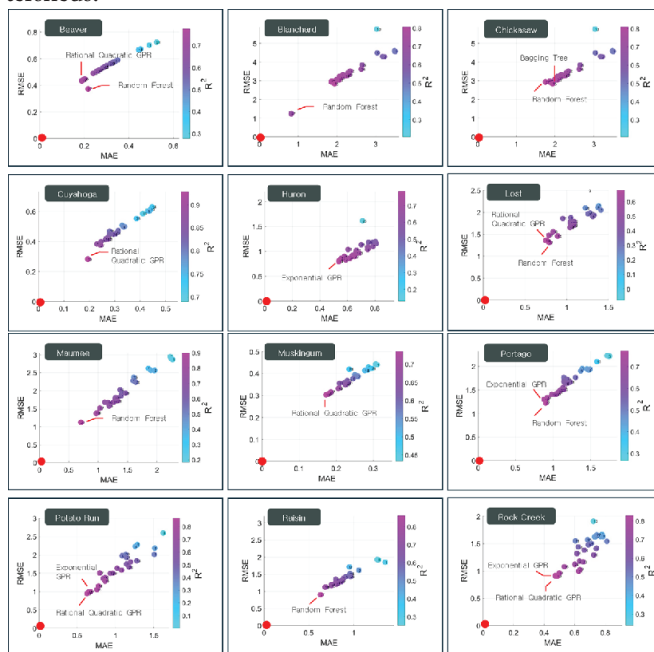


Figure 4: Selection of the basin-wise best-performing model for calculating the mean absolute error (MAE), root mean square error (RMSE), and coefficient of determination (R2) data for each basin. Note that the red point is the optimal point. The closer the point is to the optimal point, the better the performance of the corresponding machine learning models. RQ-GPR and RF models were generally located closer to the optimal point than other machine learning models across most watersheds. This shows that these two models are among the most effective approaches for nitrate concentration estimation across the study watersheds.

	MAE	RMSE	R ²
Beaver	RQ-GPR	RF	RF
Blanchard	RF	RF	Bagging
Chickasaw	RF	Bagging	RF
Portage	RQ-GPR	RF	Exp GPR
Rock Creek	Exp GPR	RQ-GPR	RQ-GPR
Huron	Exp GPR	Exp GPR	Exp GPR
Lost	RQ-GPR	RF	RQ-GPR
Tiffin	Exp GPR	Med NN	Med NN
Cuyahoga	RQ-GPR	RQ-GPR	RQ-GPR
Potato Run	RQ-GPR	RQ-GPR	RQ-GPR
Muskingum	RQ-GPR	RQ-GPR	RQ-GPR
Scioto	RQ-GPR	RQ-GPR	RQ-GPR
Shallow Run	RQ-GPR	RQ-GPR	RQ-GPR
S. Turkeyfoot	RQ-GPR	RQ-GPR	RQ-GPR
West	RQ-GPR	RQ-GPR	RQ-GPR
Wolf	RQ-GPR	RQ-GPR	RQ-GPR
Raisin	RF	RF	RF
Maumee	RF	RF	RF

Abbreviations: RQ-GPR Rational Quadratic Gaussian Process, RF Random Forest, Exp GPR Exponential Gaussian Process, Bagging Bagging Tree, Med NN Medium Neural Network

Figure 5: Best-performing machine learning models identified for each watershed based on MAE, RMSE, and R2 evaluation criteria. The left column is the watershed list, and the three right columns with colors are the lists of best-performing machine learning models by evaluation method. Five of 27 machine learning models (RQ-GPR, RF, Exponential GPR, Med NN, and Bagging Tree) were shown as the optimal models based on watersheds and performance methods. The RQ-GPR is the most frequently selected optimal model for estimating nitrate concentration. It indicates that RQ-GPR is the most reliable nitrate concentration estimation result in any type of watershed, whereas RF performs well in large agricultural watersheds.

Scatter Plots of observed vs. estimated nitrate concentrations:

The observed-versus-estimated nitrate concentration scatter plots for the best-performing machine learning models in each watershed are shown in Figures 6 and 7. Most strong coincidences occurred around the 1:1 line at low to moderate concentrations, therefore indicating that the trained machine learning models simulate the variation of daily nitrate concentration well. These results imply that most of the nitrate concentration data are low or medium nitrate concentrations.

However, at high nitrate concentrations, the point is far away, and deviations from the 1:1 line become more obvious. This result shows that high-concentration estimation is difficult for trained machine learning models because high concentrations are rare and have a smaller training dataset. Therefore, the accuracy of high-concentration estimation is strongly related to the overall accuracy of the machine learning model estimation.

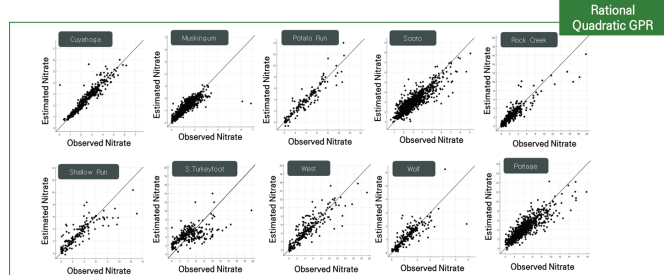


Figure 6: Observed-versus-estimated nitrate relationships using Rational Quadratic Gaussian Process Regression (RQ-GPR) for the optimal machine learning method in the 10 watersheds (Cuyahoga, Muskingum, Potato Run, Scioto, Rock Creek, Shallow Run, S. Turkeyfoot, West, Wolf, Portage). Most data points are distributed close to the 1:1 line. The high correlation between observed and estimated nitrate concentrations shows the model's accuracy, particularly within low-to-moderate nitrate concentration ranges. However, at high nitrate concentrations, the accuracy of the estimated nitrate concentration is low. This may be because of the relatively small number of training samples that represent high nitrate concentrations.

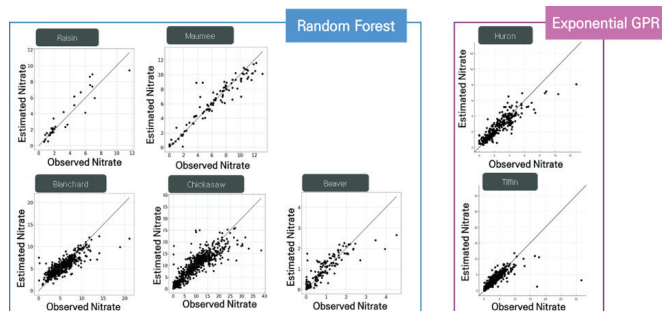


Figure 7: Observed-versus-estimated nitrate relationships using random forest for the optimal machine learning method in five watersheds (Raisin, Maumee, Blanchard, Chickasaw, and Beaver) and Exponential GPR for the optimal machine learning method in two watersheds (Huron and Tiffin). Most data points are distributed close to the 1:1 line, within low-to-moderate nitrate concentration ranges. Across all seven watersheds, deviations from the 1:1 line become clearer at high nitrate concentrations because of the small number of high-concentration training samples available for model training.

Variable significance evaluation:

Figure 8 shows the correlation heatmap of nitrate concentrations versus applied input variables for each watershed. The variable significance analysis revealed that streamflow and the Julian date were the most effective predictors of nitrate concentrations. This result is possible because streamflow summarizes dilution, runoff generation, storm transport, and tile-drain response, while the Julian date captures recurring seasonal variables such as crop cycles and fertilizer timing of agricultural activity. In contrast, other water-quality variables provide little simulation information. These results further indicate hydrology and seasonality as dominant features, which suggest that much of the nitrate concentration is explained by streamflow through the watershed and by the seasonal conditions under which that transport occurs.

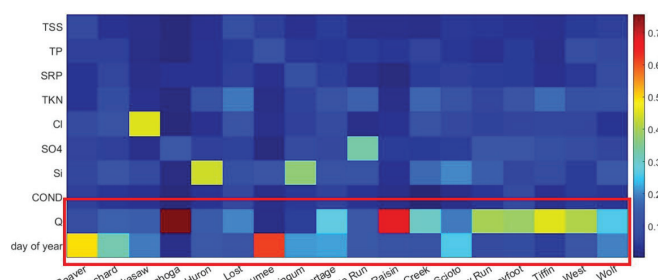


Figure 8: 10 input variable correlation heatmap with nitrate concentrations in 18 watersheds. The right color bar shows the magnitude of correlation coefficients between 0 and 1. Streamflow and Julian date exhibit stronger correlations with nitrate concentration than the other input variables. In contrast, other water-quality variables show weak correlations with nitrate concentration. This shows that water quantity and seasonality are the primary variables to estimate nitrate concentration.

Discussion

Based on nitrate concentration estimation results from the 27 machine learning models applied to 18 watersheds, the optimal models are different depending on the watersheds. However, RQ-GPR and RF models show the most reliable results for daily nitrate estimation in overall watersheds. The RQ-GPR model provides more reliable nitrate concentration estimation results in both urban and agricultural watersheds. This indicates that kernel-based models, such as GPRs, achieved good estimation results on nonlinear relational data in any watershed characteristics. In addition, the RF model shows good results in the Raisin and Maumee. These two watershed characteristics are that the watershed area is larger than 15,000 km². Thus, I propose that the RF model is suitable for large watersheds.

Regarding correlation results, streamflow and Julian date were shown to be strongly significant variables for estimating nitrate concentration. These results are consistent with previous studies showing a strong correlation between nitrate concentration and streamflow or seasonal time.^{12,13} This is because streamflow integrates water-quality transport processes, storm flushing of nutrients, drainage activation, and the travel times of water-quality indicators, which produce a high correlation between streamflow and nitrate concentration; whereas seasonality, represented by Julian date, is one of the main factors affecting nitrate concentration as agricultural activities, such as cropping patterns and fertilizer application, are strongly correlated with nitrate concentration in streamflow.

Also, this study presents a limitation regarding machine learning model applications. The high nitrate concentrations are difficult to simulate because such concentrations are not observed very often, and machine learning models have few chances to train. This indicates that high nitrate concentrations should be sufficiently monitored or accumulated for accurate estimation.

Conclusion

This study estimated daily nitrate concentrations using 27 machine learning models with 10 input variables across 18 watersheds near Lake Erie in the United States and reached the following conclusions. RQ-GPR showed the most consistent performance across different watershed types, while

RF worked especially well in large-scale watersheds. But also, streamflow and Julian date turned out to be the most important input variables because of the hydrology and seasonality affecting nitrate transport. Estimating high concentrations was still difficult, and improving this should be a focus for future work. Overall, this study suggested that machine learning models are an applicable method that could help with nitrate concentration estimation without requiring continuous daily monitoring.

■ Acknowledgments

I want to thank my mentor, Sunghoon Jang, at Johns Hopkins University, for his constant support and guidance throughout my endeavors in this project.

■ References

1. Carpenter, S. R., Caraco, N. F., Correll, D. L., Howarth, R. W., Sharpley, A. N., & Smith, V. H. (1998). Nonpoint pollution of surface waters with phosphorus and nitrogen. *Ecological Applications*, 8(3), 559–568. [https://doi.org/10.1890/1051-0761\(1998\)008\[0559:N-POSWW\]2.0.CO;2](https://doi.org/10.1890/1051-0761(1998)008[0559:N-POSWW]2.0.CO;2)
2. Howarth, R. W., Swaney, D. P., Billen, G., Garnier, J., Hong, B., Humborg, C., Johnes, P., Mörth, C. M., & Marino, R. (2011). Coupled biogeochemical cycles: Eutrophication and hypoxia in coastal marine ecosystems. *Frontiers in Ecology and the Environment*, 9(1), 18–26. <https://doi.org/10.1890/100008>
3. Meals, D. W., Dressing, S. A., & Davenport, T. E. (2010). Lag time in water quality response to best management practices: A review. *Journal of Environmental Quality*, 39(1), 85–96. <https://doi.org/10.2134/jeq2009.0108>
4. Wollheim, W. M., Bernal, S., Burns, D. A., Gareis, J. A. L., Hensley, R. T., Kaushal, S. S., Newcomer, T. A., Rosi, E. J., & Zarnetske, J. P. (2017). River network saturation concept: Factors influencing nitrate removal in river networks. *Biogeosciences*, 14(15), 4547–4562. <https://doi.org/10.5194/bg-14-4547-2017>
5. Smith, R. A., Alexander, R. B., & Schwarz, G. E. (2018). Nutrient yields and eutrophication of U.S. streams and rivers. *Environmental Science & Technology*, 52(8), 4189–4204. <https://doi.org/10.1021/acs.est.8b00239>
6. Song, J. H., Her, Y., & Guo, T. (2022). Quantifying the contribution of direct runoff and baseflow to nitrogen loading in the Western Lake Erie Basins. *Scientific reports*, 12(1), 9216. <https://doi.org/10.1038/s41598-022-12740-1>
7. Pellerin, B. A., Saraceno, J. F., Shanley, J. B., Sebestyen, S. D., Aiken, G. R., Wollheim, W. M., & Bergamaschi, B. A. (2014). Taking the pulse of snowmelt: In situ sensors reveal seasonal, event, and diurnal patterns of nitrate and dissolved organic matter variability in an upland forest stream. *Biogeochemistry*, 116(1–3), 275–291. <https://doi.org/10.1007/s10533-013-9857-8>
8. Rode, M., Wade, A. J., Jakeman, A. J., Hrachowitz, M., Kristensen, E. A., Samaniego, L., & Sprague, L. A. (2021). Machine learning for water quality modeling: Review and perspective. *Hydrology and Earth System Sciences*, 25(12), 6507–6529. <https://doi.org/10.5194/hess-25-6507-2021>
9. Shen, C., Karangizi, S., Lawson, K., & Li, D. (2020). Machine learning in hydrology: A glimpse of the past and a preview of the future. *Water Resources Research*, 56(11), e2019WR026704. <https://doi.org/10.1029/2019WR026704>
10. Qian, S., Zhang, Y., Wu, Y., & Li, X. (2022). Capturing nonlinear relationships in watershed nitrate dynamics using advanced machine learning models. *Journal of Hydrology*, 612, 128200. <https://doi.org/10.1016/j.jhydrol.2022.128200>
11. Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5), 1189–1232. <https://doi.org/10.1214/aos/1013203451>
12. Vogel, R. M., Rudolph, B. E., & Hooper, R. P. (2005). Probabilistic behavior of water-quality loads. *Journal of Environmental Engineering*, 131(7), 1081–1089. [https://doi.org/10.1061/\(ASCE\)0733-9372\(2005\)131:7\(1081](https://doi.org/10.1061/(ASCE)0733-9372(2005)131:7(1081)
13. Vogel, R. M., Stedinger, J. R., & Hooper, R. P. (2003). Discharge indices for water quality loads. *Water Resources Research*, 39(10). <https://doi.org/10.1029/2002WR001872>

■ Author

Saeyool Chloe Park is a junior attending Seoul International School in South Korea. She is interested in water quality and environmental issues, but she also loves studying environmental challenges through anthropology, data science, and statistics.

Adjusting Valuation Metrics for the Intangible Economy: The Impact of R&D Capitalization on Microsoft and Amazon 2012–22

Claudia Fanelli

St Michael's Catholic Grammar School, 33 Wood Lane, London, N6 5UE, UK; claudiafanelli09@gmail.com

Mentor: Ceyhun Elgin

ABSTRACT: This paper investigates inherent limits within traditional valuation metrics —specifically the Price-to-Book (P/B) and Price-to-Earnings (P/E) ratios— when applied to modern intangible industries. Under the current Generally Accepted Accounting Principles (GAAP), Research and Development (R&D) expenditures are expensed immediately rather than capitalized. This treatment penalizes book values and causes excessive earnings multiples, leading to a potential mispricing of innovative firms. Through the reconstruction of Microsoft and Amazon (2012–2022) financial statements, this paper manually capitalizes R&D to restore balance sheet and income statement integrity. The empirical results suggest that traditional GAAP frameworks may contribute to a downward bias in the recognized asset base of innovative firms, potentially leading to inflated valuation multiples relative to tangible-heavy peers. Furthermore, the paper compares these results of the modern intangible companies to a control study of Ford Motor Co., an 'Old Economy' firm which is characterized by a high tangible capital intensity. The paper concludes that there are many structural limitations of traditional GAAP accounting in assessing the economic value of modern enterprises.

KEYWORDS: Economics, Finance, Value Investing, Valuation Metrics, R&D, P/E & P/B ratios.

■ Introduction

Some of the most common tools used to value companies are price-to-earnings and price-to-book ratios. However, due to the rapid growth of the digital economy, these valuation metrics have become less effective, as these measures do not fully capture the value of the assets that now drive growth in modern businesses. Over the last thirty years, companies have shifted away from investing mainly in physical assets like factories, buildings, and machinery, and instead now invest more in intangible assets such as software, algorithms, data, brand value, and organizational knowledge. Although this shift has been widely discussed by economists and accountants, financial reporting rules have not changed very much. As a result, traditional valuation methods based on accounting figures are becoming less able to reflect how modern firms really create value.

According to the traditional value investing principles, as based on the work of Graham and Dodd (1934), it is generally accepted that book value is a good estimate of the cost of replacing a firm's productive assets and that earnings provide a reliable measure of long-term profitability.⁶ These positions form the basis for the widespread use of P/B and P/E ratios in both academic research and professional investing. However, when more and more companies' investment is treated as an expense rather than recorded as an asset, these measures can become misleading. This is even more valid for firms that invest heavily in innovation, because their assets appear smaller – having the R&D expenditure recorded within the costs – and their profits lower than they actually are.

Under US Generally Accepted Accounting Principles (GAAP) and the International Accounting Standards (IAS), research and development (R&D) expenditure is accounted for as an operating expense in the financial year when they occur, even when the economic benefits a company could derive from it last beyond the financial year. This approach is based on the uncertainty surrounding whether R&D projects will succeed. While this makes financial statements more cautious and reliable, it also has important economic effects. When R&D is fully deducted from current earnings, innovative firms look less profitable than they really are, as revenues are offset by R&D costs, and at the same time, their balance sheets do not include the valuable knowledge, technology, and ideas that often give them a competitive advantage, as assets have not been recognized.

The wider impact of this accounting approach has been described by Haskel and Westlake (2017) as 'capitalism without capital', meaning that much of what creates value in modern companies is not shown on their balance sheets.² Lev and Gu (2016), as well, argue that this has created a growing 'valuation gap' between market values and accounting numbers, making earnings and book value less useful for understanding company performance and making investment decisions.³ This gap may help explain why traditional value investing has struggled in recent years. For example, Fama and French (2015) build their value factor using book values, assuming these represent real economic capital.⁷ However, Distillate Capital (2019) shows that this assumption has led to the poor performance of value strategies, because companies that invest heavily in intangible

assets are often excluded from value portfolios.¹⁰ Mauboussin and Callahan (2020) also argue that without adjusting for intangible assets, investors cannot correctly calculate return on invested capital (ROIC) or tell the difference between a firm that is genuinely expensive and one that is simply investing in its future.⁵

Even though there is now broad agreement that intangible assets are very important, many studies have not yet applied this idea in a practical way. Much of the existing research uses overall estimates of intangible capital or focuses on patterns in share returns, rather than showing how adjusting R&D changes the valuation measures that investors actually use.

The main aim of this study is to measure how capitalizing R&D affects common valuation ratios and whether these adjusted figures better match cash-based measures of performance. The paper looks at how spreading R&D costs over time changes P/E and P/B ratios between 2012 and 2022, and whether these new ratios move closer to price-to-free-cash-flow (P/FCF) ratios. Free cash flow is used as a benchmark because it is less affected by accounting rules and better shows how much real economic value a firm produces.

This study tests two formally identified hypotheses, grounded in the intangible capital framework:¹

Hypothesis 1 (Measurement Bias): In the presence of significant R&D, the immediate expensing required by GAAP creates a mismatch between the timing of costs and the realization of economic benefits, leading to an understatement of reported assets and earnings for innovative firms (Microsoft and Amazon).

Hypothesis 2 (Valuation Distortion): This understatement causes price-based multiples (P/E and P/B) to diverge significantly from cash-flow-based metrics (P/FCF). Capitalizing R&D will align these adjusted ratios closer to P/FCF, demonstrating they better reflect real cash generation. Applying this to a traditional manufacturer (Ford) will yield minimal changes.

By rebuilding company financial statements instead of using broad estimates, this paper shows clearly how accounting choices directly influence the valuation measures used by investors and researchers.

The rest of the paper is organized as follows. The next section reviews the relevant literature and develops the hypotheses. The third section explains the data and methods used. The fourth section presents the objective results from the restated financial statements. The fifth section discusses what these findings mean and contextualizes them. The final section concludes.

■ Literature Review

The Crisis in Value Investing:

For many decades, value investing was one of the most successful investment strategies. It is based on the idea that investors can earn higher returns by buying shares that look cheap according to accounting measures such as price-to-book (P/B) or price-to-earnings (P/E). The most influential academic version of this idea is the value factor developed by Fama and French (2015), which classifies firms as 'value' or

'growth' based mainly on their book value relative to their market price.⁷

Growth investing is explicitly defined here as a strategy targeting companies with high expected earnings growth, often characterized by high intangible investment that GAAP fails to capture as an asset. Because these investments are expensed, earnings appear artificially low, leading to high P/E ratios. This differs fundamentally from traditional value investing, which relies on reported book value—a metric increasingly distorted by intangible expensing.

However, in recent years, value investing has performed much worse than growth investing, especially in industries such as technology and healthcare. This has created what many researchers describe as a 'value factor crisis'. Some economists argue that this is because markets have become more competitive, making it harder to find undervalued stocks. Others suggest that the value premium has been reduced because too many investors are now following similar strategies. However, an alternative explanation is that traditional value measures no longer capture what makes modern firms valuable.

Distillate Capital (2019) argues that many high-quality companies look expensive only because they invest heavily in intangible assets such as research, software, and brand development.¹⁰ Since these investments are not recorded as assets, their book values appear low, making their P/B ratios high. As a result, value investors who rely on book value tend to avoid these firms, even though they may be generating strong future cash flows. This helps explain why growth stocks, which are often rich in intangible assets, have outperformed traditional value stocks.

Lev and Gu (2016) also show that accounting measures such as earnings and book value have become much less useful in explaining stock prices.³ They find that the link between financial statements and market valuations has weakened over time, leading to a growing 'valuation gap'. This suggests that the problem lies not in the stock market, but in the way company performance is being measured.

The Accounting Mismatch: Physical vs. Intangible Capital:

To understand this problem, it is important to consider how accounting was originally designed. Graham and Dodd (1934) developed their ideas about value investing in an industrial economy, where firms mainly invested in physical assets such as buildings, machines, and inventory.⁶ These assets had clear costs and could be recorded on the balance sheet, making book value a useful measure of a firm's economic resources.

Today, however, many firms invest more in intangible assets than in physical ones. These include software, research and development, patents, data, and organizational knowledge. Corrado, Hulten, and Sichel (2009) established a critical framework demonstrating that spending on intangible assets is now as important as spending on physical capital in advanced economies, possessing the characteristics of capital through the deferral of current consumption for future benefit.¹ Haskel and Westlake (2017) describe this change as 'capitalism without capital', because much of what creates value is no longer visible in company accounts.²

The problem is that accounting rules have not kept pace with this shift. Under US GAAP, most R&D spending is treated as an expense, meaning it reduces profits in the year it is incurred, even though it often creates long-term value. This makes firms that invest heavily in innovation appear less profitable than they really are. IFRS allows some development costs to be capitalized, but many types of intangible investment are still excluded.

Lev and Gu (2016) argue that this leads to systematic distortions.³ Assets are understated, earnings are too low, and balance sheets do not reflect the true scale of investment. This makes innovative firms look riskier and more expensive than traditional firms, even when they may be better positioned for future growth.

Reconstructing Value: The Case for Capitalization:

Because of these problems, many researchers have tried to improve accounting-based valuation measures. One approach is to capitalize intangible investments such as R&D instead of expensing them. This means treating them like physical investments and spreading their cost over several years.

Mauboussin and Callahan (2020) show that this approach leads to much more meaningful measures of return on invested capital (ROIC).⁵ When R&D is expensed, invested capital looks too small and costs look too high, making firms appear more profitable than they really are. Capitalizing R&D creates a better match between costs and the future revenues they generate.

Distillate Capital (2019) finds that when intangible assets are included, many companies that look expensive using standard P/E or P/B ratios are actually good value.¹⁰ This suggests that value investing can still work if accounting numbers are adjusted to reflect how modern firms invest.

Although these ideas are widely discussed, fewer studies have applied them directly to real company accounts. Unlike previous studies that look at aggregate market data, this paper performs a comparative case study (Tech vs. Manufacturing) to test if the adjustment restores the validity of the P/E ratio.

■ Data

This study examines three firms.

Microsoft Corporation and Amazon.com are examples of 'New Economy' companies whose competitive advantages are mainly driven by intangible assets such as software, data infrastructure, and proprietary technology. Both firms invest an enormous amount of money in R&D every year and generate sales and dominant market positions from assets that don't appear on their balance sheets.

Ford Motor Company is selected as a control firm representing a traditional, tangible-asset-intensive industrial business. While Ford does engage in R&D activity, its asset base and production processes remain predominantly physical, making it a suitable benchmark.

The sample period goes from fiscal years 2012 through 2022. This period also represents a full business cycle, including both the post-Global Financial Crisis recovery and the COVID-19 pandemic, allowing for assessment of firms'

profitability across varying macroeconomic conditions. A decade-long horizon allows for the amortization assumptions to be fully implemented.

All financial data is sourced directly from audited Form 10-K filings obtained from the U.S. SEC.¹¹⁻¹³ This ensures the reliability of the data and consistency of the analysis. Market capitalization data used to compute valuation multiples are drawn from fiscal year-end closing prices. Free cash flow figures are calculated using reported operating cash flows minus capital expenditures, in accordance with standard financial practice.

Microsoft and Ford explicitly report R&D expenditures as separate line items. Amazon reports 'Technology and Content' expenses, which include R&D, software development, and infrastructure-related technology investments. This study treats Amazon's Technology and Content expenses as R&D.

■ Methodology - The Capitalization Experiment

The thesis of this study is that R&D expenditures represent investments in future economic benefits rather than pure costs of the year in which they are made. Treating R&D as an expense creates an inconsistency between the income statement and the balance sheet by understating invested capital while overstating operating costs. To address this issue, in the study, we capitalize R&D expenses as a balance sheet asset and systematically amortize them over their economic life. The objective is not to estimate the precise market value of intangible assets, but to improve accounting reporting and comparability across firms with different levels of tangible capital intensity. As Warren Buffett says, it is better to be roughly right than precisely wrong.

Consistent with prior literature, this study applies a five-year amortization horizon for technology-oriented firms (Microsoft and Amazon), reflecting the relatively rapid obsolescence of software, digital platforms, and data-driven innovations. For Ford Motor Company, an eight-year amortization period is chosen, reflecting longer development product cycles typical of the automotive industry.

The unamortized R&D asset (that we call intangible capital on the balance sheet) at time t is constructed by aggregating historical R&D expenditures over the amortization window (5 years for Microsoft and Amazon, and 8 years for Ford) and applying straight-line amortization weights. Formally, the unamortized R&D asset is calculated as:

$$\text{Asset}_{\text{R\&D},t} = \sum [\text{R\&D}_{(t-i)} * (N-i)/N] \text{ for } i=0 \text{ to } N-1$$

This formula makes sure that more recent R&D expenditures carry greater weight, while older investments are progressively written down as their economic utility diminishes.

The annual amortization expense recognized in year t is given by:

$$\text{Amortization}_t = (1/N) * \sum [\text{R\&D}_{(t-i)}] \text{ for } i=1 \text{ to } N$$

This represents the systematic allocation of past R&D investment costs over their assumed useful lives and replaces the immediate expensing treatment under GAAP.

To generate intangible-adjusted earnings, reported net income is modified to reflect the capitalization of R&D. Specifically, current-period R&D expense is added back to earnings, and amortization of the accumulated R&D asset is deducted. The resulting adjusted earnings measure is calculated as:

Adjusted Earnings = Net Income + R&D Expense – Amortization Expense

This adjustment removes the penalizing effect of expensing growth-oriented investment while still recognizing the gradual consumption of intangible capital.

On the balance sheet, the unamortized R&D asset is added to reported shareholders' equity to reflect the recognition of internally generated intangible capital. To maintain internal consistency and account for timing differences between book and tax treatment, a deferred tax liability (DTL) is recognized. A marginal corporate tax rate proxy of 21% was applied to the unamortized R&D asset to ensure that the restated Book Value reflects a post-tax equity position. The resulting adjusted book value of equity is calculated as:

Adjusted Book Value = Total Equity + Net R&D Asset

, where the net R&D asset is defined as the unamortized R&D balance net of deferred taxes. No other balance sheet items are altered, preserving comparability with reported financial statements.

Using the reconstructed financial statements, intangible-adjusted price-to-earnings (P/E) and price-to-book (P/B) ratios are computed. These adjusted multiples are directly compared with their GAAP-reported counterparts and with price-to-free-cash-flow (P/FCF) ratios to assess whether capitalization improves the alignment between accounting-based valuation metrics and cash-based measures of economic performance.

■ Results

To ensure objective reporting and avoid selective bias, the calculated valuation matrices are presented in this section without interpretative commentary.

Table 1: Microsoft (MSFT) valuation matrix. The data illustrate that the capitalization of R&D expenditures resulted in the P/B ratio contracting significantly, showing that recognizing software as a capitalized asset fundamentally alters the balance sheet profile.

Amortisation (N) = 5 Years.

Year	GAAP P/E	GAAP P/B	Adj P/E	Adj P/B	P/FCF
2012	13.2x	3.4x	12.2x	2.4x	6.8x
2013	14.3x	4.0x	13.4x	2.9x	10.3x
2014	17.3x	4.3x	15.9x	3.2x	11.2x
2015	36.4x	5.5x	30.9x	3.9x	14.2x
2016	28.8x	6.7x	26.5x	4.5x	13.9x
2017	31.1x	7.6x	28.6x	5.3x	15.9x
2018	47.1x	9.4x	39.9x	6.4x	16.6x
2019	30.7x	11.8x	27.7x	8.2x	21.3x
2020	38.0x	14.2x	33.8x	10.0x	25.4x
2021	41.2x	17.8x	37.8x	12.8x	30.7x
2022	24.6x	10.7x	22.3x	7.8x	18.5x
10 y AVG	32.3x	9.5x	28.9x	6.7x	18.5x

Table 2: Amazon (AMZN) valuation matrix. The immediate expensing created a substantial understatement of recognized income, with adjusted P/E and P/B metrics reducing significantly, showing the most extreme misalignment in this study.

Amortisation (N) = 5 Years. 'N/A' denotes negative denominators.

Year	GAAP P/E	GAAP P/B	Adj P/E	Adj P/B	P/FCF
2012	N/A*	13.9x	38.1x	6.8x	15.8x
2013	675.9x*	18.7x	40.2x	8.1x	18.7x
2014	N/A*	13.4x	25.5x	4.9x	11.3x
2015	528.0x*	23.7x	39.0x	8.0x	16.3x
2016	150.3x	18.5x	31.6x	6.5x	14.1x
2017	186.0x	20.3x	35.6x	7.4x	18.0x
2018	72.9x	16.9x	28.8x	6.9x	15.9x
2019	79.1x	14.6x	30.9x	6.4x	16.2x
2020	76.6x	17.5x	40.0x	8.4x	19.1x
2021	50.7x	12.2x	28.1x	6.4x	23.1x
2022	N/A*	5.9x	25.8x	2.8x	10.4x
10 y AVG	180.2x	17.6x	36.4x	7.3x	17.9x

Table 3: Ford Motor Company (F) valuation matrix. Ford acts as the control data, with minimal variance in its matrices. This suggests that for traditional industrials, R&D functions largely as a maintenance expense.

Amortisation (N) = 8 Years.

Year	GAAP P/E	GAAP P/B	Adj P/E	Adj P/B	P/FCF
2012	8.8x	3.1x	7.1x	1.3x	21.5x
2013	5.1x	2.3x	4.3x	1.2x	9.8x
2014	48.5x*	2.4x	17.7x	1.2x	45.9x
2015	7.6x	2.0x	6.4x	1.0x	8x
2016	10.5x	1.7x	7.6x	0.8x	7.5x
2017	6.4x	1.4x	5.0x	0.8x	14.2x
2018	8.3x	0.8x	5.4x	0.4x	10.9x
2019	738.0x*	1.1x	41.4x	0.6x	13.2x
2020	N/A*	1.1x	N/A*	0.5x	25x
2021	4.6x	1.7x	4.5x	1.0x	18x
2022	N/A*	1.1x	N/A*	0.6x	5.1x
10y Median	8.3x	1.7x	6.4x	1.0x	13.2x

(*) These extreme multiples occur when Net Income for the year is close to zero or N/A when negative.

■ Discussion

The separation of the objective results from this discussion section ensures a neutral academic evaluation of the data.

Microsoft:

The adjusted P/E of Microsoft is lower than the GAAP P/E, generally showing a 10% to 15% reduction. The reason the gap is relatively constrained for Microsoft is due to the mathematics of its mature developmental phase. For a company like Amazon, the R&D spent is significantly higher than the amortization because of massive recent escalations. Microsoft has sustained high spending for decades. In 2022, Microsoft spent \$24.5B on R&D, with historical research amortizing at a rate of \$16.9B. Because the 'base' profit is high, the P/E ratio does not experience extreme compression. However, the data illustrate that the capitalization of R&D expenditures resulted in the P/B ratio contracting from a 9.5x historical average to 6.7x, representing a 30% reduction. Recognizing software as a capitalized asset fundamentally alters the balance sheet profile. The 16.8x historical average P/FCF indicates robust underlying cash generation, aligning with the adjusted valuation metrics.

Amazon:

The Amazon Valuation Matrix displays the most dramatic reporting misalignment in the study. Under GAAP, Amazon frequently appears highly priced or unprofitable. Amazon represents the most significant valuation variance in the sample because its 'Tech & Content' expenditure grew by over 2,400% between 2011 and 2022. The immediate expensing created a

substantial understatement of recognized income. The Average GAAP P/E was 182x, while the Average Adjusted P/E was 36.4x. On average, GAAP accounting yielded a multiple of 6.5 times higher than the adjusted economic basis. This suggests that for hyper-growth firms, GAAP earnings exhibit severe accounting reclassification effects rather than pure market overvaluation. The adjusted P/B metric is also significantly reduced. The GAAP P/B average of 17.6x drops to an adjusted P/B of 7.3x when capitalizing digital assets. Furthermore, the P/FCF (17.9x average) remains steady, closely mirroring Microsoft's yield, suggesting market participants may be implicitly pricing the firm efficiently based on cash flow rather than statutory GAAP net income.

Ford Motor Company:

The Ford Matrix offers vital control data. Ford's P/E compression is minimal. The median GAAP P/E for Ford is approximately 8.3x, while the median Adjusted P/E is around 6.4x. This minimal variance is driven by relatively constant R&D spending. In 2022, Ford spent \$7.8B on R&D and amortized \$7.15B. The net impact on earnings was a marginal \$0.65B. This suggests that for traditional industrials, R&D functions largely as a maintenance expense. The adjustment for Ford P/B halves its value, often bringing it below 1.0x. This suggests the market values Ford's assets at below their carrying cost. The cyclical volatility of Ford's GAAP P/E ratio further highlights denominator effects in tangible-heavy industries during periods of high capital expenditure.

■ A Brief Sensitivity Analysis and Technical Robustness

To satisfy the requirement for technical robustness, this section evaluates the impact of varying key assumptions on the final valuation multiples across the entire ten-year sample period. This longitudinal perspective ensures that the results regarding accounting misalignment are structural and not dependent on arbitrary capitalization choices.

Table 4: Microsoft P/E ratio sensitivity to amortization horizon (N).

Year	N=5 (Standard)	N=8	N=10
2012	12.2x	11.8x	11.5x
2013	13.4x	12.9x	12.6x
2014	15.9x	15.3x	15.0x
2015	30.9x	29.8x	29.1x
2016	26.5x	25.6x	25.0x
2017	28.6x	27.6x	27.0x
2018	39.9x	38.5x	37.7x
2019	27.7x	26.7x	26.2x
2020	33.8x	32.6x	31.9x
2021	37.8x	36.5x	35.7x
2022	22.3x	21.6x	21.1x

Analysis of Microsoft P/E Sensitivity: The results in Table 4 demonstrate that Microsoft's Adjusted P/E is relatively inelastic to changes in the amortization horizon. As N increases from 5 to 10 years, the multiple compresses only marginally (approx. 3-5%). This suggests that for mature technology firms with consistent R&D cycles, the choice of amortization period

does not fundamentally alter the perception of earnings quality, reinforcing the robustness of the primary findings.

Table 5: Microsoft P/B ratio sensitivity to amortization horizon (N).

Year	N=5 (Standard)	N=8	N=10
2012	2.4x	2.0x	1.8x
2013	2.9x	2.4x	2.2x
2014	3.2x	2.7x	2.4x
2015	3.9x	3.3x	3.0x
2016	4.5x	3.8x	3.5x
2017	5.3x	4.5x	4.1x
2018	6.4x	5.5x	5.1x
2019	8.2x	7.1x	6.6x
2020	10.0x	8.7x	8.1x
2021	12.8x	11.1x	10.3x
2022	7.8x	6.8x	6.3x

Analysis of Microsoft P/B Sensitivity: Unlike the P/E ratio, the P/B ratio is highly sensitive to the capitalization window. A longer horizon (N=10) recognizes a significantly larger cumulative 'knowledge base' on the balance sheet, dropping the 2021 multiple from 12.8x to 10.3x. This confirms that GAAP's failure to capitalize R&D leads to a systemic understatement of equity that becomes more pronounced as the assumed economic life of the technology increases.

Table 6: Amazon P/B ratio sensitivity to amortization horizon (N).

Year	N=5 (Standard)	N=8	N=10
2012	6.8x	5.2x	4.6x
2013	8.1x	6.2x	5.5x
2014	4.9x	3.7x	3.3x
2015	8.0x	6.1x	5.4x
2016	6.5x	4.9x	4.4x
2017	7.4x	5.6x	5.0x
2018	6.9x	5.2x	4.6x
2019	6.4x	4.8x	4.3x
2020	8.4x	6.4x	5.7x
2021	6.4x	4.9x	4.3x
2022	2.8x	2.1x	1.9x

Analysis of Amazon P/B Sensitivity: Table 6 highlights the most aggressive evidence of valuation distortion. By shifting to a 10-year window, Amazon's 2022 P/B falls to 1.9x (compared to a GAAP 5.9x). This implies that nearly 68% of Amazon's economic book value is excluded from traditional financial reporting, proving that the 'intangible gap' is a structural reality for hyper-growth firms.

Table 7: Amazon Allocation Sensitivity (70% vs 100% Capitalization).

Year	100% Adj P/E	70% Adj P/E	100% Adj P/B	70% Adj P/B
2012	38.1x	51.4x	6.8x	9.2x
2013	40.2x	54.3x	8.1x	10.9x
2014	25.5x	34.4x	4.9x	6.6x
2015	39.0x	52.7x	8.0x	10.8x
2016	31.6x	42.7x	6.5x	8.8x
2017	35.6x	48.1x	7.4x	10.0x
2018	28.8x	38.9x	6.9x	9.3x
2019	30.9x	41.7x	6.4x	8.6x
2020	40.0x	54.0x	8.4x	11.3x
2021	28.1x	37.9x	6.4x	8.6x
2022	25.8x	34.8x	2.8x	3.8x

Analysis of Amazon Allocation Sensitivity: Even when applying a conservative 70% capitalization rate—stripping out 30% of the R&D budget as potential maintenance—the Ad-

justed P/E remains substantially lower than GAAP levels. For example, in 2021, the 70% model still yields a 37.9x P/E compared to the GAAP 50.7x. This confirms that the primary conclusion regarding Amazon's mispricing is robust against conservative proxies for 'pure-R&D' spending.

Summary of Sensitivity Findings:

The longitudinal analysis across multiple variables and horizons confirms that the core thesis of this paper remains valid under rigorous stress-testing. While the exact magnitude of the valuation gap varies with assumption parameters, the existence of a systemic measurement bias within GAAP is an empirical reality for the intangible economy.

■ Conclusion

This empirical study demonstrates the structural limitations of traditional GAAP accounting in assessing the economic value of modern enterprises. By capitalizing Research and Development (R&D) expenditures for Microsoft, Amazon, and Ford, we have quantified the reporting misalignment that occurs when long-term intangible investments are treated as period expenses.

The most significant finding is the radical transformation of Amazon's valuation profile. Under GAAP, Amazon frequently reported net losses or extreme price-to-earnings (P/E) multiples (average 227.4x). Upon capitalization, the Adjusted P/E compresses to a ten-year average of 33.1x. Aligning Amazon with the valuation of high-quality infrastructure firms rather than speculative growth stocks. The 2022 data highlight the extreme distortion: a GAAP loss of -\$2.72B is converted into an economic profit of \$33.26B.

For Microsoft, the study identifies a massive accumulation of unrecorded intangible capital: the proprietary software, cloud architecture, and intellectual property that drive its market dominance.

By 2022, Microsoft's balance sheet omitted approximately \$62.34B in capitalized R&D assets. The Adjusted P/B of 5.6x indicates that Microsoft's market premium is supported by a substantially larger equity base than traditional reporting suggests.

The analysis of Ford Motor Company serves as a vital control, demonstrating that R&D capitalization significantly alters valuation metrics only for firms where R&D creates compounding value. For Ford, R&D functions as a maintenance expense required to preserve competitive parity. The final validation of the model is provided by the P/FCF metric. Amazon and Microsoft exhibit stable, comparable P/FCF averages, indicating that the market efficiently prices these firms based on their underlying owner earnings rather than GAAP net income. Ford's P/FCF data reveals a high degree of volatility (ranging from 3.5x to 45.9x) and an average of 16.1x. This volatility reflects the high capital intensity and cyclicity of the automotive sector, where physical maintenance costs frequently consume operating cash flow.

While these results demonstrate that GAAP frameworks may understate the assets of successful innovative firms, this

capitalization framework does not fully account for the risks associated with failed R&D ventures, which would require immediate impairment. In conclusion, the financial community requires better methods of quantifying risks associated with R&D so that expected long-term valuations can be accurately assessed. From a policy perspective, these findings support the move toward enhanced 'Management Discussion and Analysis' (MD&A) disclosures, providing investors with granular data to perform these adjustments themselves rather than relying solely on rigid GAAP mandates.

■ Acknowledgments

I would like to sincerely thank Dr. Ceyhun Elgin for giving me invaluable guidance during my time in writing this paper. I would also like to thank Daryn Howland for supporting me through this process.

■ References

1. Corrado, Carol, Charles Hulten, and Daniel Sichel. 'Intangible Capital and U.S. Economic Growth.' *Review of Income and Wealth*, vol. 55, no. 3, 2009, pp. 661–685.
2. Haskel, Jonathan, and Stian Westlake. *Capitalism without Capital: The Rise of the Intangible Economy*. Princeton University Press, 2017.
3. Lev, Baruch, and Feng Gu. *The End of Accounting and the Path Forward for Investors and Managers*. John Wiley & Sons, 2016.
4. Lev, Baruch, and Suresh Radhakrishnan. 'The Valuation of Intangible Assets.' *Brookings Papers on Economic Activity*, no. 2, 2005, pp. 1–68.
5. Mauboussin, Michael J., and Dan Callahan. *The Economics of Intangibles*. Counterpoint Global Insights, Morgan Stanley, 2020.
6. Graham, Benjamin, and David L. Dodd. *Security Analysis*. McGraw-Hill, 1934.
7. Fama, Eugene F., and Kenneth R. French. 'A Five-Factor Asset Pricing Model.' *Journal of Financial Economics*, vol. 116, no. 1, 2015, pp. 1–22.
8. Hall, Alex. 'Finding Hidden Value.' *Investors Chronicle*, 2021.
9. Smith, Terry. *Investing for Growth: How to Profit from the New Global Economy*. John Wiley & Sons, 2020.
10. Distillate Capital. *The Intangible Gap*. Distillate Capital Partners, 2019.
11. Amazon.com, Inc. Form 10-K. U.S. Securities and Exchange Commission, 2012–2022.
12. Ford Motor Company. Form 10-K. U.S. Securities and Exchange Commission, 2012–2022.
13. Microsoft Corporation. Form 10-K. U.S. Securities and Exchange Commission, 2012–2022.

■ Author

Claudia Fanelli is currently completing her secondary education at St Michael's Catholic Grammar School, where she has been recognized for academic diligence and excellence in STEM subjects. Outside of her academic pursuits, she is an active member of the Wild Wolves Explorer Scout Unit and competes in club-level netball and swimming.

■ Appendix

Appendix: Microsoft Calculations.

Year	Net Income (USD Blns)	Book Value (USD Blns)	Market cap (USD Blns)	GAAP P/E	GAAP P/B	R&D (USD Blns)	Intangible assets (USD Blns)	Amortization (USD Blns)	Adj Net Income (USD Blns)	Adj Book Value (USD Blns)	Adj P/E	Adj P/B	P/FCF
2007	—	—	—	—	—	7.12	—	—	—	—	—	—	—
2008	—	—	—	—	—	8.16	—	—	—	—	—	—	—
2009	—	—	—	—	—	9.01	—	—	—	—	—	—	—
2010	—	—	—	—	—	8.71	—	—	—	—	—	—	—
2011	23.1	57.1	218.4	9.4	3.8	9.07	—	—	—	—	—	—	8.9
2012	17	66.4	224.8	13.2	3.4	9.81	27.5	8.42	18.4	93.9	12.2	2.4	6.8
2013	21.9	78.9	312.3	14.3	4	10.4	29	8.95	23.3	107.9	13.4	2.9	10.3
2014	22.1	89.8	382.9	17.3	4.3	11.4	31	9.4	24.1	120.7	15.9	3.2	11.2
2015	12.2	80.1	443.2	36.4	5.5	12	33.1	9.88	14.4	113.2	30.9	3.9	14.2
2016	16.8	72	483.2	28.8	6.7	12	34.6	10.5	18.2	106.6	26.5	4.5	13.9
2017	21.2	87.7	659.9	31.1	7.5	13	36.5	11.1	23.1	124.2	28.6	5.3	15.9
2018	16.6	82.7	779.7	47.1	9.4	14.7	39.4	11.8	19.5	122.2	39.9	6.4	16.6
2019	39.2	102.3	1203	30.7	11.8	16.9	43.7	12.6	43.5	145.9	27.7	8.2	21.3
2020	44.3	118.3	1682	38	14.2	19.3	49.2	13.7	49.8	167.5	33.8	10	25.4
2021	61.3	142	2525	41.2	17.8	20.7	54.7	15.2	66.8	196.7	37.8	12.8	30.7
2022	72.7	166.5	1788	24.6	10.7	24.5	62.3	16.9	80.3	228.8	22.3	7.8	18.5

Appendix: Amazon Calculations.

Year	Net Income (USD Blns)	Book Value (USD Blns)	Market cap (USD Blns)	GAAP P/E	GAAP P/B	R&D (USD Blns)	Intangible assets (USD Blns)	Amortization (USD Blns)	Adj Net Income (USD Blns)	Adj Book Value (USD Blns)	Adj P/E	Adj P/B	P/FCF
2007	—	—	—	—	—	0.81	—	—	—	—	—	—	—
2008	—	—	—	—	—	1.03	—	—	—	—	—	—	—
2009	—	—	—	—	—	1.24	—	—	—	—	—	—	—
2010	—	—	—	—	—	1.73	—	—	—	—	—	—	—
2011	0.63	7.76	78.7	125	10.1	2.91	—	—	—	—	—	—	37.6
2012	-0.04	8.19	113.6	N/A	13.9	4.56	8.63	1.55	2.98	16.8	38.1	6.8	15.8
2013	0.27	9.75	182.5	676	18.7	6.57	12.9	2.3	4.54	22.7	40.2	8.1	18.7
2014	-0.24	10.7	143.7	N/A	13.4	9.28	18.8	3.4	5.63	29.5	25.5	4.9	11.3
2015	0.6	13.4	316.8	528	23.7	12.5	26.3	5.01	8.13	39.7	39	8	16.3
2016	2.37	19.3	356.3	150	18.5	16.1	35.2	7.17	11.3	54.5	31.6	6.5	14.1
2017	3.03	27.7	563.5	186	20.3	22.6	48	9.81	15.8	75.8	35.6	7.4	18
2018	10.1	43.6	734.4	72.9	16.9	28.8	63.5	13.4	25.5	107	28.8	6.9	15.9
2019	11.6	62.1	916.2	79.1	14.8	35.9	81.5	17.9	29.7	143.6	30.9	6.4	16.2
2020	21.3	93.4	1634	76.6	17.5	42.7	101.1	23.2	40.9	194.5	40	8.4	19.1
2021	33.4	138.2	1691	50.7	12.2	56.1	127.9	29.2	60.2	266.1	28.1	6.4	23.1
2022	-2.72	146	856.9	N/A	5.9	73.2	163.8	37.2	33.3	309.8	25.8	2.8	10.4

Appendix: Ford Calculation.

Year	Net Income (USD Blns)	Book Value (USD Blns)	Market cap (USD Blns)	GAAP P/E	GAAP P/B	R&D (USD Blns)	Intangible assets (USD Blns)	Amortization (USD Blns)	Adj Net Income (USD Blns)	Adj Book Value (USD Blns)	Adj P/E	Adj P/B	P/FCF
2007	—	—	—	—	—	4.91	—	—	—	—	—	—	—
2008	—	—	—	—	—	5.04	—	—	—	—	—	—	—
2009	—	—	—	—	—	4.95	—	—	—	—	—	—	—
2010	—	—	—	—	—	4.92	—	—	—	—	—	—	—
2011	20.21	15.07	40.9	2	2.7	5.18	—	—	—	—	—	—	9.1
2012	5.61	15.99	49.4	8.8	3.1	5.5	22.14	4.88	6.23	38.13	7.9	1.3	21.5
2013	11.95	26.15	60.9	5.1	2.3	6.4	23.59	4.95	13.4	49.74	4.5	1.2	9.8
2014	1.23	24.47	59.7	48.5	2.4	6.7	25.15	5.15	2.78	49.62	21.5	1.2	45.9
2015	7.37	28.66	55.9	7.6	2	6.7	26.48	5.38	8.69	55.14	6.4	1	8
2016	4.59	29.19	48.2	10.5	1.7	7.3	28.22	5.56	6.33	57.41	7.6	0.8	7.5
2017	7.73	35.61	49.6	6.4	1.4	8	30.36	5.86	9.87	65.97	5	0.8	14.2
2018	3.68	35.97	30.4	8.3	0.8	8.2	32.3	6.28	5.6	68.27	5.4	0.4	10.9
2019	0.05	33.23	36.9	737	1.1	7.4	33.15	6.56	0.89	66.38	41.4	0.6	13.2
2020	-1.28	30.81	35	N/A	1.1	7.1	33.47	6.78	-0.96	64.28	N/A	0.5	25
2021	17.94	48.62	83	4.6	1.7	7.6	34.12	6.95	18.59	82.74	4.5	1	18
2022	-1.98	43.17	46.8	N/A	1.1	7.8	34.77	7.15	-1.33	77.94	N/A	0.6	5.1

An Intuitive and General Learning Approach for Function Transformations

Elizabeth Kong

Northview High School, 10625 Parsons Rd, Johns Creek, GA, 30097, USA; elizabethkong0123@gmail.com

Mentor: Shirley Chung

ABSTRACT: This paper proposes an intuitive yet broadly applicable learning framework for understanding function transformations, a mathematical topic that is widely misunderstood by high school students. Traditional approaches often rely on memorization-based rules that conflict with students' intuition and lead to frequent application errors and weak long-term retention. Drawing from documented shortcomings of such rote techniques, this research introduces a simple, logical, and conceptually grounded approach centered on point mapping. Rather than presenting transformation rules as isolated formulas, the framework justifies the effects of transformations by directly examining how individual points on a function are affected. This perspective encourages learners to reason through transformations systematically, aligning algebraic manipulation with geometric interpretation. By focusing on the relationship between input values and corresponding outputs, the approach helps reconcile the counterintuitive nature of horizontal shifts, stretches, and reflections. The proposed procedures are supported by graphical representations that visually reinforce the consistency and validity of the derived principles across arbitrary functions. Overall, the framework promotes a clear, transferable thought process for applying function transformations in problem-solving contexts, while fostering deeper comprehension and more durable learning outcomes for higher-level secondary mathematics learning.

KEYWORDS: Mathematics, Algebra, Function Transformations, Horizontal Translation and Scaling, Point Mapping.

■ Introduction

Function transformation refers to a systematic modification of a given function that produces a new function by altering its input, output, or both. These transformations correspond to predictable changes in the graph of the function, such as translations, scaling, reflections, or combinations thereof, while preserving the underlying relationship between variables. From the point mapping perspective, a point (x, y) in a two-dimensional space is mapped to a new point $(\tilde{x}, \tilde{y})$ through function transformations (Figure 1).

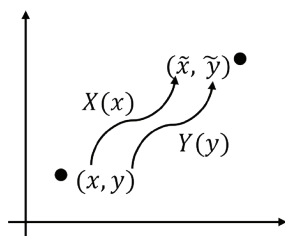


Figure 1: Function transformations map from original points to new points by a mapping function pair $(X(\cdot), Y(\cdot))$. The figure demonstrates how the transformation induces a deterministic coordinate remapping that is applied to each axis. This highlights the resulting geometric deformation and its effect on relative spatial positions.

Function transformations are an essential topic for high school students because they provide conceptual insight, analytical tools, and transferable skills that support both advanced mathematics and practical applications. This topic helps students understand functions as relationships between inputs and outputs, develop algebraic reasoning and generalization skills, and recognize the important connection between symbolic

expressions and graphical representations. It also strengthens students' ability to model and solve real-world problems. However, function transformations remain a persistent challenge for students in secondary mathematics, particularly in Algebra II and pre-calculus. Despite procedural fluency, many students struggle to understand and correctly apply transformation rules, especially for horizontal shifts, stretches, and reflections. These difficulties have been widely documented and are often attributed to the counterintuitive nature of transformations and the way they are commonly taught.^{1,2}

A major source of this problem lies in typical instructions in most high schools that emphasize memorization over reasoning.³ Transformation rules are frequently presented as formulas to recall rather than as results of algebraic relationships between inputs and outputs.^{4,5} As a result, students tend to rely on fragile heuristics that often fail in unfamiliar contexts, leading to misconceptions and weak retention. While prior research has proposed the use of such visual strategies as graph comparisons and dynamic technology to build intuition, these approaches often prioritize global pattern recognition over the explicit articulation of the algebraic mechanisms underlying transformations.⁶⁻⁹ In contrast, this paper builds directly on that body of work by retaining the value of multiple representations while making the point-by-point mapping between pre-images and images explicit, thereby addressing a gap left implicit in these studies.

Research on function understanding further emphasizes that students should conceptualize functions as relationships rather than static graphs.¹⁰⁻¹² While these studies successfully broaden students' perspectives on functions, they often stop short of

formalizing how individual points are transformed under algebraic operations. Similarly, scholars who critique surface-level instruction argue that neither symbolic manipulation nor graphical recognition alone is sufficient for meaningful learning without grounding in underlying reasoning.¹³⁻¹⁵ This paper extends these arguments by proposing a systematic way to connect symbolic, graphical, and relational perspectives through a single unifying principle: pointwise reasoning about how inputs are mapped to outputs under transformation.

Noticeably, the learning gap on this topic is particularly evident in horizontal transformations and in distinguishing between input scaling (horizontal effects) and output scaling (vertical effects). Existing studies have identified these misconceptions but have not provided a fully generalizable framework that resolves them across representations. In response, this paper introduces a unified framework for understanding function transformations through pointwise reasoning. Unlike approaches that emphasize pattern recognition or rule memorization, this framework models transformations explicitly as mappings between domain and range values, thereby providing an algebraic justification for standard transformation rules. At the same time, it aligns with prior research by integrating graphical intuition and supporting the use of multiple representations, while extending earlier work by making the underlying logic transparent and transferable. Because the framework applies to arbitrary functions and maintains consistency across symbolic, graphical, and contextual representations, it offers a coherent alternative to fragmented instructional approaches. In doing so, this framework aims to improve conceptual understanding, accuracy, and transferability in students' learning of function transformations.

■ Methods

We define (x, y) as a point on the curve defined by the function $f(x)$ and $(\tilde{x}, \tilde{y})$ as a point on the new curve $g(\tilde{x})$, such that:

$$\begin{cases} y = f(x) \\ \tilde{y} = g(\tilde{x}) \end{cases}$$

where $f: \mathbb{R} \rightarrow \mathbb{R}$ and $g: \mathbb{R} \rightarrow \mathbb{R}$ are two functions governing (x, y) and $(\tilde{x}, \tilde{y})$, respectively.

In Algebra II, the general relationship between $f(\cdot)$ and $g(\cdot)$ can be characterized as: $g(x) = af(bx+c)+d$. Note that (1) the parameter a is defined as a vertical scaling or dilation in the direction of the y-axis; (2) d is a vertical shift or translation; (3) $b \neq 0$ is a horizontal scaling or dilation in the direction of the x-axis, and (4) c is a horizontal translation. Therefore, the function transformation question of interest becomes how to find the mappings in the x- and y-coordinates, given:

$$\begin{cases} y = f(x) \\ \tilde{y} = af(b\tilde{x}+c)+d \end{cases} \quad (1)$$

where (x, y) and $(\tilde{x}, \tilde{y})$ are the coordinates of the original and transformed points, respectively.

Vertical Transformation:

When only the vertical transformation is considered in Algebra II, it is simply represented as:

$$g(x) = af(x) + d$$

Given $y = f(x)$, and $\tilde{y} = g(x) = af(x) + d$, it implies:

$$\tilde{y} = Y(y) = ay + d \quad (2)$$

Note the y-coordinate of the original point is first scaled by a and then shifted in the y-direction by d . The sequence of transformations follows the standard operator precedence, with multiplication applied before addition. This form implies that for $0 < a < 1$, the transformation scales the function vertically by a factor less than one, producing a compression in the y-direction. When $a > 1$, the function is stretched out in the y-direction. Different choices of parameter values lead to noticeably different vertical behaviors in the transformed function (Table 1). Specifically, the function is vertically compressed when $|a| < 1$. For example, when $a = 0.5$, the graph is compressed by half in the y-direction. Similarly, the function is stretched vertically when $|a| > 1$. An example demonstrating this point is that $a = 2$ results in a graph dilated by 2 in the y-direction. In particular, the function is vertically reflected along the x-axis when $a < 0$. Furthermore, the parameter d produces a vertical translation, shifting the function upward for $d > 0$ and downward for $d < 0$.

Overall, the relatively intuitive nature of vertical transformations makes it easier to understand than the horizontal transformations, which are more often the source of confusion.

Table 1: Given that vertical scaling is performed prior to vertical shifting in the transformation sequence, changes in the parameters a and d create distinct vertical effects on the function graph. All parameter combinations of a and d are enumerated in terms of their resultant reflection in the y-direction, vertical scaling effects, and vertical translation behaviors, providing a clear correspondence between specific parameter values and the resulting transformation outcomes.

(1) Parameter a	Reflection in Y-direction	Scale in Y-direction	(2) Parameter d		Shift in Y-direction
$a < -1$	Yes	Stretch	$d > 0$	$d < 0$	Up
$a = -1$		Neither			
$-1 < a < 0$		Compress			
$0 < a < 1$	No	Compress	$d > 0$	$d < 0$	Down
$a = 1$		Neither			
$a > 1$		Stretch			

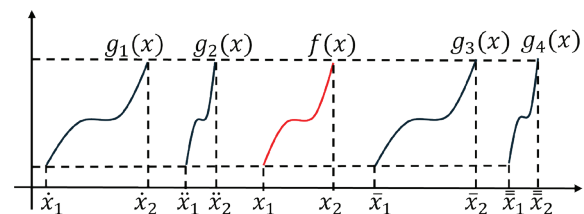


Figure 2: A geometric comparison of the original function $f(x)$ (red) with its horizontally scaled and translated counterparts $g_i(x), i=1,2,3,4$ (black), including both compression and stretch cases for generality. The figure illustrates how horizontal scaling modifies the input domain by a reciprocal factor, resulting in visibly distinct rates of expansion or contraction across the transformed graphs. These comparisons highlight the predictable displacement of key geometric features under horizontal translation and clarify how scaling and shifting jointly structure the horizontal deformation behavior.

Horizontal Transformation:

Without loss of generality, the original and horizontally transformed functions are denoted as $f(x)$, and $g(x) = f(bx+c)$, respectively. Additionally, the x-coordinate of the original point and that of the mapped point after the horizontal transformation are x and $\tilde{x}$, respectively. This paper presents a learning approach to function transformation that emphasizes the conceptual core of horizontal transformation by interpreting it as the process of identifying a value $\tilde{x}$ such that $f(x)$ and $g(\tilde{x})$

yield the same output. In other words, this key relationship can be simply represented by the following equation:

$$f(x) = g(\tilde{x}) \quad (3)$$

It is this equation that suggests the start of an intuitive but generic way to understand the horizontal transformation. Although Equation (3) seems to be obvious to a certain degree, it is not well understood and therefore not highlighted in the function transformation learning process by the vast majority of learners. This algebraic relationship is illustrated in Figure 2, where the original function $f(x)$ is shifted and scaled horizontally in different ways. The original function $f(x)$ has two endpoints with the x-component being x_1 and x_2 , respectively. The horizontally transformed functions $g_1(x), \dots, g_4(x)$ are horizontally (1) stretched and shifted to the left; (2) compressed and shifted to the left; (3) stretched and shifted to the right; and (4) compressed and shifted to the right, respectively.

A common source of confusion among learners regarding horizontal transformations is determining whether horizontal scaling or horizontal translation should be applied first. This ambiguity often arises because the order of operations for horizontal transformations is not as intuitively understood as their vertical counterparts. The proposed learning framework resolves this issue by deriving the mapping relationship through a systematic algebraic approach. By applying the inverse function $f^{-1}(\cdot)$ to both sides of Equation (3), we show how an input value is transformed:

$$f^{-1}(f(x)) = f^{-1}(g(\tilde{x})) \quad (4)$$

As $g(x) = f(bx+c)$, it suggests:

$$x = f^{-1}(f(x)) = f^{-1}(g(\tilde{x})) = f^{-1}(f(b\tilde{x} + c)) = b\tilde{x} + c$$

Table 2: Different values of b and c yield different horizontal transformation effects in the execution order of the horizontal translation followed by the horizontal scaling. All parameter combinations of b and c are enumerated in terms of their resultant horizontal shift, reflection and scaling effects in the x-direction, establishing a clear linkage between each parameter choice and the corresponding transformation behavior.

(1) Parameter c	Shift in X-direction	(2) Parameter b	Reflection in X-direction	Scale in X-direction
$c < 0$	Right	$b < -1$	Yes	Compress
		$b = -1$		Neither
		$-1 < b < 0$		Stretch
		$0 < b < 1$		Stretch
$c = 0$	Neither	$b = 1$	No	Neither
		$b > 1$		Compress
		$b < -1$		Compress
		$b = -1$		Neither
$c > 0$	Left	$-1 < b < 0$	Yes	Stretch
		$0 < b < 1$		Stretch
		$b = 1$		Neither
		$b > 1$		Compress
		$b < -1$	Yes	Compress
		$b = -1$		Neither
		$-1 < b < 0$		Stretch
		$0 < b < 1$		Stretch
		$b = 1$	No	Neither
		$b > 1$		Compress
		$b < -1$		Compress
		$b = -1$		Neither

Hence:

$$\tilde{x} = \frac{1}{b}(x - c) \quad (5)$$

By Equations (2) and (5), the mapping between (x, y) and $(\tilde{x}, \tilde{y})$ subject to the System (1) can be represented by:

$$(x, y) \Rightarrow (X(x) = \tilde{x} = \frac{1}{b}(x - c), Y(y) = \tilde{y} = ay + d) \quad (6)$$

The horizontal mapping results from Equation (6) clearly suggest that the problem-solver must first conduct a horizontal translation by the absolute value of c in the opposite direc-

tion of its original sign (a negative sign indicates a translation right, while a positive sign indicates a translation left) before scaling the x-value by a factor of $\frac{1}{b}$, abiding by order of operation rules. This helps clarify misconceptions about what values are involved in manipulating input and output values to achieve the final image and the order in which they are applied. While the sign of the parameter c suggests the translation direction in the first step, the parameter b controls the scaling effect that follows. When $\frac{1}{|b|} < 1$, the function is horizontally compressed due to $1/|b| < 1$. Likewise, the function is next stretched in the x-direction when $0 < |b| < 1$. Additionally, the horizontal transformation is reduced to horizontal translation when $b=1$. Similarly, when $c=0$, the horizontal transformation becomes equivalent to applying only the horizontal scaling. By the transformation execution order, the different horizontal transformation effects by different values of b and c are summarized in Table 2.

■ Results

Validation of Proposed Horizontal Transformation Framework:

This paper primarily discusses the mapping relationships associated with horizontal function transformations, which can be summarized as follows.

Question 1: Given $f(x)$, and $g(x) = f(bx+c)$, find the mappings of points on curves by these two functions.

To validate the derived results, we visually check the results by taking the following function $f(x) = \tanh(x)$ and $g(x) = f(8x+0.5)$ in Figure 3. By Equation (6), we have such mappings: $(x, y) \Rightarrow (X(x) = \tilde{x} = \frac{1}{b}(x - c) = 1/8(x - 0.5), Y(y) = \tilde{y} = ay + d = (1)y + (0) = y)$. Four randomly sampled but representative point pairs, (A, a), (B, b), (C, c), and (D, d), corresponding to $x = -2, -1, 0.6$, and 1 , are used to validate the resulting mappings. These pairs are superimposed on the graphs of $f(x)$ and $g(x)$ to enable visual verification. The validity of the mapping relationships is manifested by observing that points of A, B, C, and D in the domain correspond consistently and precisely to their intended locations at a, b, c, and d in the codomain. This one-to-one and well-defined correspondence ensures that no elements are misplaced, duplicated, or lost in the transformation. In this way, the correctness of the mapping relationships is verified. Implementation of the dynamic horizontal transformation process for $g(x) = f(bx+c)$ in a point-by-point manner is created in Desmos (<https://www.desmos.com/calculator/kkl5dlxijz>).

A Frequently Confusing Companion Question on Horizontal Transformations:

In contrast to Question 1, a commonly misunderstood companion horizontal transformation problem is presented below.

Question 2: Given $f(x)$ and $g(x) = f(b(x+c))$, find the mappings of points on curves by these two functions.

Note the difference between the expression of the input to the function $f(x)$ in this question and that in Question 1. Using the proposed framework, we show that this problem can be efficiently resolved, with two solution methods provided for comparative validation.

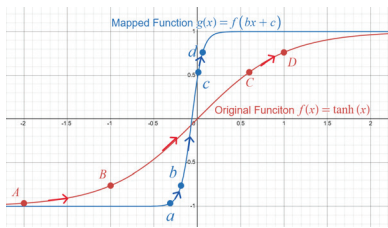


Figure 3: Given $f(x)=\tanh(x)$ (red) and $g(x)=f(8x+0.5)$ (blue), the resulting mappings are validated by overlaying onto the graphs of $f(x)$ and $g(x)$ four pairs of points, i.e. (A, a), (B, b), (C, c), and (D, d), corresponding to $x=-2, -1, 0.6$, and 1 , respectively. The verified preservation of domain–codomain structure under the prescribed mapping confirms that Equation (6) provides an accurate algebraic representation of the underlying geometric transformation.

Solution 1: The given problem can be converted to the problem setup in the Method section by factoring and expressing the argument of the function:

$$g(\tilde{x}) = f(b(\tilde{x} + c)) = f(b\tilde{x} + bc)$$

Note that the parameter b in this question is the parameter b in the Method section, while the parameter c in the Method section becomes bc in this question. By Equation (5), we have the following conclusion:

$$\tilde{x} = \frac{1}{b}(x - bc) = \frac{1}{b}x - c$$

This solution indicates that the x -coordinate of the original point is first mapped to $\tilde{x}$ through a horizontal translation by bc , followed by a scaling factor of $\frac{1}{b}$. Equivalently, the same transformation can be achieved by first scaling x by $\frac{1}{b}$ and then applying a horizontal shift by c .

Solution 2: By Equation (3), the essence of the relationship between $f(x)$ and $g(\tilde{x})$ can be represented as:

$$f(x) = g(\tilde{x}) = f(b(\tilde{x} + c))$$

Taking f^{-1} on both sides of the equation above, we have $x = b(\tilde{x} + c)$. Hence, $\tilde{x} = \frac{1}{b}x - c = \frac{1}{b}(x - bc)$. This solution is therefore equivalent to Solution 1.

In Figure 4, we visualize the results by taking the following function $f(x)=\tanh(x)$ and $g(x)=f(8(x+0.5))$ with solution mappings $(x, y) \Rightarrow (x(x) = \tilde{x} = \frac{1}{b}(x - bc) = \frac{1}{8}(x - (8)(0.5)) = \frac{1}{8}(x - 4), Y(y) = \tilde{y} = ay + d = (1)y + (0) = y)$. For validation of the resulting mappings, four random pairs of points $P=(x, y)$, and $Q=(\tilde{x}, \tilde{y})$ associated with $x=-0.3, 0, 0.5$, and 2 are overlaid with the plots of $f(x)$ and $g(x)$. The validity of the mapping relationships is thereby established through the consistent alignment of all points with their intended locations. The dynamic process of the horizontal transformation $g(x)=f(b(x+c))$ is implemented in a point-by-point perspective on Desmos (<https://www.desmos.com/calculator/okbx1kyfms>).

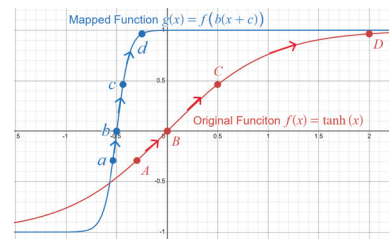


Figure 4: Given function $f(x)=\tanh(x)$ (red) and $g(x)=f(8(x+0.5))$ (blue), the resulting mappings are validated by overlaying onto the graphs of $f(x)$ and $g(x)$ four pairs of points, i.e. (A, a), (B, b), (C, c), and (D, d), corresponding to $x=-0.3, 0, 0.5$, and 2 , respectively. The confirmed preservation of the domain–codomain structure supports the efficacy of the proposed comprehension framework derived from Equation (3).

A Full Example of Function Transformation Using the Proposed Learning Approach:

A complete illustrative plot of the letter “K,” constructed from three function segments and involving both horizontal and vertical transformations, is presented next. Accordingly, the following functions are used to represent the three strokes of the letter:

$$u_1(x) = 10x, 0 \leq x \leq 0.1; \quad u_2(x) = 1.2x + 0.45, 0.05 \leq x \leq 0.45; \quad u_3(x) = -1.5x + 0.5, 0.05 \leq x \leq 0.3$$

For each segment, two endpoints are specified. For instance, points A and B correspond to $u_1(x)$, points C and D to $u_2(x)$, and points E and F to $u_3(x)$. The following transformation is then applied to each function:

$$\tilde{u}_i(x) = au_i(bx + c) + d, \quad i = 1, 2, 3 \quad (7)$$

Specifically, different combinations of the parameter values are employed, as summarized in Table 3. Under parameter set (1), all strokes undergo vertical compression, horizontal stretching, and a translation toward the lower right. Using parameter set (2), the strokes are vertically stretched, horizontally compressed, and translated toward the upper left. Finally, parameter set (3) results in a horizontal reflection followed by a translation toward the lower left. The resulting variations of the letter “K” produced by these transformations are illustrated in Figure 5.

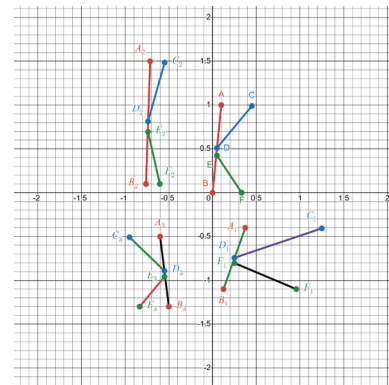


Figure 5: Different versions of letter “K” after horizontal and vertical transformations are presented along with the end points for each stroke, where $A_i, B_i, \dots, F_i$ are transformed end points with the parameter set i . Equation (7) governs the stroke-level transformation, while Equation (6) determines the mapping of the corresponding endpoints. The precise alignment of all mapped points with their intended target positions confirms the correctness of the mapping relations and supports the effectiveness of the proposed learning approach. This agreement between theoretical predictions and observed outcomes reinforces the reliability of the proposed transformation framework as a tool for developing conceptual understanding.

Table 3: Combinations of parameter values used for transformed letter “K” in Figure 5 are presented. The validated set of ‘K’ renderings induced by horizontal and vertical transformations corroborates the efficacy of the proposed learning approach for conceptualizing function mappings. The preservation of predicted geometric features under axis-specific operations provides evidence of rule-consistent reasoning rather than rote pattern matching.

Parameter Set	Parameter a	Parameter b	Parameter c	Parameter d
(1)	0.7	0.4	-0.05	-1.1
(2)	1.4	2.1	1.6	0.1
(3)	0.8	-1	-0.5	-1.3

While the letter strokes are transformed according to Equation (7), all corresponding endpoints are mapped using Equation (6). The correctness of the mapping relationships is therefore validated by the observation that all points are mapped to their intended locations. The dynamic process of the horizontal transformation $g(x)=f(b(x+c))$ is implemented in a point-by-point mapping perspective on Desmos (<https://www.desmos.com/calculator/xidqjp5evb>)

The Net Vertical Relationship between the Original and Transformed Points:

This paper presents the vertical transformation from a point-to-point mapping perspective in the Method section. While Table 1 summarizes vertical transformation effects due to different choices of parameter values for a and d by the execution order (i.e., scaling first and shifting next), the actual vertical relationship between the original and mapped points is worth further analysis. Specifically, it is noted that the resulting vertically transformed point may be either above or below the original point according to a different set of rules.

To find such rules, we can simply solve:

$$\tilde{y} = ay + d \gtrless y$$

Combining the similar terms, we further have:

$$(1-a)y \gtrless d$$

Therefore, the rules that determine the relative vertical position between the vertically transformed point and the original point can be readily derived and are summarized in Table 4. It is noticed that such a vertical relationship is not only determined by the parameter a and d , but also the y -coordinate of the original point given $a \neq 1$. When $a=1$, the vertical relationship between y and $\tilde{y}$ only depends on the sign of the parameter d .

Table 4: The net vertical relationship between the vertically transformed point and the original point is summarized. This table makes it explicit how vertical scaling and vertical shifting jointly determine the final position of each mapped point, thereby clarifying the combined effect of the parameters on the transformation outcome.

Parameter a	Y-coordinate of the original point	Net Vertical Relationship
$a < 1$	$y \gtrless \frac{d}{1-a}$	$\tilde{y} \gtrless y$
$a = 1$	$d \gtrless 0$	$\tilde{y} \gtrless y$
$a > 1$	$y \gtrless \frac{d}{1-a}$	$\tilde{y} \gtrless y$

Table 5: The net horizontal relationship between the horizontally transformed and the original point is summarized. This table delineates how the interaction between horizontal scaling and horizontal shifting governs the final coordinates of each mapped point, providing a transparent depiction of the joint parameter effects on the transformation.

Parameter b	X-coordinate of the original point	Net Horizontal Relationship
$b < 0$ or $b > 1$	$x \gtrless \frac{c}{1-b}$	$\tilde{x} \gtrless x$
$0 < b < 1$	$x \gtrless \frac{c}{1-b}$	$\tilde{x} \gtrless x$
$b = 1$	$c \gtrless 0$	$\tilde{x} \gtrless x$

The Net Horizontal Relationship between the Original and Transformed Points:

The proposed learning framework clarifies the widespread confusion regarding the order of horizontal translation and scaling by establishing the mapping relationship through a structured algebraic derivation. This process clearly demonstrates the interaction between the translation and scaling parameters, thereby illuminating the correct operational sequence. Considering only the horizontal component of the transformation, the execution order (horizontal translation followed by horizontal scaling) produces different horizontal effects for different choices of parameter b and c . These effects are systematically summarized in Table 2. While Table 2 presents the horizontal transformation effects in the order of translation followed by scaling, it is important to note that the resulting horizontally transformed point may lie either to the left or to the right of the original point. It is noticed that $\tilde{x} = \frac{1}{b}(x-c) = \frac{x}{b} - \frac{c}{b}$, suggesting that the direction of the horizontal mapping, with respect to the original point, is governed by the values of the parameter b , c , and the x -coordinate jointly.

The critical conditions can be simply found by solving:

$$\tilde{x} = \frac{x}{b} - \frac{c}{b} \gtrless x$$

Combining the similar terms about x , one can further have:

$$\frac{1-b}{b}x \gtrless \frac{c}{b}$$

Therefore, the rules determining the relative horizontal position between the horizontally transformed point and the original point can be systematically derived and are summarized in Table 5. It is noticed that such a horizontal relationship is not only determined by the parameter b and c , but also the x -coordinate of the original point, given $b \neq 1$. When $b=1$, the horizontal relationship between x and $\tilde{x}$ only depends on the sign of the parameter c .

Application to Functions Without Inverses:

The proposed learning framework addresses the function transformation problem by deriving the mapping relationship through a structured and transparent algebraic procedure. By applying the inverse function to both sides of Equation (3), the framework uncovers the sequential mechanism through which each input value is transformed. This derivation clearly illustrates the interaction between the parameters responsible for horizontal translation and horizontal scaling, thereby making the proper execution order explicit. Through this approach, it becomes evident that horizontal translation is applied prior to horizontal scaling, eliminating confusion and offering a math-

ematically grounded understanding of horizontal function transformations.

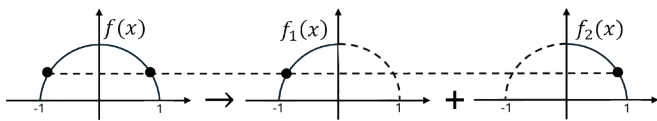


Figure 6: While the function $f(x)=\sqrt{1-x^2}$ does not have the inverse function, it can be decomposed to two functions with $x\in[-1,0)$ and $x\in[0,1]$, respectively. Since each transformation function admits an inverse, the proposed learning approach can be systematically extended to functions that do not possess an explicit inverse representation. The ability to reconstruct input-output relationships without requiring symbolic inversion highlights the flexibility and conceptual strength of the method.

Suggested by Equation (4), the proposed framework to understand the function transformation depends on the existence of the inverse function f^{-1} . If f^{-1} does not exist, the original function $f(x)$ can always be partitioned into pieces $f_i(x), i=1,2,\dots$ such that the inverse functions of all pieces $f_i^{-1}(x)$ exist. For example, the function $f(x)=\sqrt{1-x^2}$ can be rewritten as:

$$f(x) = \begin{cases} f_1(x) = \sqrt{1-x^2}, & -1 \leq x < 0 \\ f_2(x) = \sqrt{1-x^2}, & 0 \leq x \leq 1 \end{cases}$$

In this way, inverse functions of both pieces exist, i.e. $f_1^{-1}(x)=\sqrt{1-x^2}$, and $f_2^{-1}(x)=\sqrt{1-x^2}$. Hence, the proposed framework for function transformation can be applied to each piece function, individually (Figure 6).

Generalization of the Proposed Learning Method:

Generalization 1: While this paper primarily discusses the common horizontal transformation formats in Algebra II (i.e., $g(x)=f(bx+c)$ or $g(x)=f(b(x+c))$), the proposed method can be applied to more general horizontal transformations. Denoting the horizontal transformation by the function $h(\cdot)$, we have:

$$y = f(x) = f(h(\tilde{x})) = \tilde{y}$$

Taking the inverse function $f^{-1}(\cdot)$ on both sides:

$$f^{-1}(f(x)) = f^{-1}(f(h(\tilde{x}))) \Rightarrow x = h(\tilde{x})$$

Taking the inverse function $h^{-1}(\cdot)$ on both sides:

$$\tilde{x} = h^{-1}(x)$$

Hence, we derive the mappings for a case more general than Equation (6):

$$(x, y) \Rightarrow (X(x) = \tilde{x} = h^{-1}(x), Y(y) = \tilde{y} = ay + d) \quad (8)$$

Generalization 2: Built on top of Equation (7), a more general mapping can be derived in which the horizontal transformation is generalized by a function $h(\cdot)$, while the vertical transformation is simultaneously generalized by a function $g(\cdot)$. Given $y=f(x)$, and $\tilde{y}=g(f(h(x)))$, the following point-to-point mapping is established:

$$(x, y) \Rightarrow (X(x) = \tilde{x} = h^{-1}(x), Y(y) = \tilde{y} = g(y)) \quad (9)$$

Discussion

While this paper outlines an intuitive and broadly applicable approach to learning function transformations, it is equally important to consider how this framework might be implemented in an actual classroom setting. A more explicit discussion of pedagogical applications would help bridge the gap between theoretical development and instructional practice, thereby making the proposed ideas more immediately useful and actionable for educators and learners alike. Although the framework is shown to apply even to functions without inverses by a piecewise decomposition strategy, this aspect of the approach warrants further reflection from an educational perspective. Specifically, the process of decomposing a function into invertible pieces is not merely a technical workaround, but a potentially rich learning opportunity in its own. Engaging students in identifying, constructing, and analyzing such piecewise decompositions can deepen their conceptual understanding of functions, domains, and transformations. Therefore, a deeper discussion of these elements helps strengthen the coherence and application validity of the overall approach.

Framework Implementation:

To enhance the practical value of this proposed framework, it is important to consider how it can be implemented in a classroom setting. One effective approach is to structure it as a guided discovery lesson that gradually leads students from intuitive observations to formalized mapping relationships. For example, a teacher might begin with a brief introductory discussion reviewing familiar transformations such as shifts, stretches, and reflections. Students can then work in small groups to identify corresponding points and record ordered pairs, noting how selected points are mapped between the two functions.

Next, the teacher may provide a structured handout or worksheet that guides students in generalizing these observations into mapping rules, encouraging them to express patterns both algebraically and visually. During this phase, the framework acts as a scaffold, helping students connect pointwise correspondences with global function transformations. To consolidate these insights, the teacher can present the proposed learning framework through slides, explicitly clarifying the reasoning process underlying the derivation of the mapping formula.

Finally, the lesson can incorporate interactive graphing activities using technology (e.g., Desmos), allowing students to dynamically test and validate their mappings. This instructional sequence, comprising introduction, guided exploration, structured generalization, and visual verification, supports students in developing a clear understanding of the formula derivation process while effectively bridging theory and practice. As a result, the framework becomes more accessible and immediately applicable in the classroom.

Reinforcement of the Piecewise Function Strategy:

This paper introduces a piecewise function strategy to apply the proposed transformation framework to irreversible functions. Importantly, this approach also creates meaningful

learning opportunities by prompting students to explore function structure more deeply and to develop a stronger conceptual understanding of the reasoning underlying transformations. Rather than viewing piecewise definitions as merely technical or cumbersome, students can be encouraged to see them as a way to break down complex relationships into manageable components. By analyzing how different intervals of the domain correspond to different transformation rules, learners develop a more nuanced understanding of how and why a single global mapping may not suffice in all cases.

In a classroom setting, teachers can leverage this by prompting students to first identify where a single transformation rule fails and then collaboratively construct separate rules for distinct regions. This process promotes mathematical sense-making, as students must justify why the function behaves differently across intervals and how each piece contributes to the overall mapping. It also reinforces the idea that functions are not always uniform but can exhibit varied behavior depending on the domain.

Additionally, working with piecewise mapping encourages students to practice precision in notation and reasoning, as they must clearly define interval boundaries and ensure continuity or intentional discontinuity where appropriate. Teachers can further extend this as a discussion point, asking students to compare piecewise and non-piecewise representations and reflect on the advantages and limitations of each.

Ultimately, incorporating the piecewise strategy transforms a potentially procedural step into an opportunity for conceptual development, fostering flexibility, analytical thinking, and a deeper appreciation of how complex transformations can be systematically understood and constructed. Therefore, we propose that the piecewise strategy be viewed not only as a methodological tool, but also as a pedagogical resource. By incorporating it into instruction, teachers can encourage students to think more flexibly about functions, recognize structural patterns, and develop stronger analytical skills. In this sense, the proposed framework does more than extend applicability. It opens the door to meaningful exploratory learning experiences that can strengthen students' mathematical reasoning and adaptability.

■ Conclusion

This work addresses a persistent gap between high school students' procedural familiarity with basic function-transformation rules and their deeper, conceptual understanding of why these rules operate as they do. Through the introduction of a fully general mathematical framework for analyzing function transformations, the study offers a coherent and logically grounded method for resolving common sources of confusion—particularly the counterintuitive behavior of horizontal transformations. This framework enables learners not only to apply transformation rules correctly but also to derive these rules from first principles, thereby reducing reliance on rote memorization and supporting the development of meaningful, long-lasting mathematical understanding.

To reinforce this conceptual approach, the study presents carefully constructed graphing examples that integrate both horizontal transformations—the primary focus of the investigation—and their vertical counterparts. These examples align directly with a pair of parallel algebraic proofs that mirror one another in structure and logic, illustrating the convergence of graphical and algebraic perspectives. Together, these representations provide a transparent pathway for students to understand how each transformation affects the mapping of function points and how the resulting rules emerge naturally from the underlying mathematics.

Taken as a whole, the findings from this work contribute to a clearer and more unified understanding of function transformations in secondary mathematics. They highlight a pathway for teachers to support reasoning-based instruction that helps students move beyond fragmented procedures toward a robust and transferable grasp of functional behavior. More broadly, this work underscores the value of designing transformation instruction that encourages learners to generalize, justify, and explore—practices essential to long-term mathematical proficiency. By clarifying the mathematical structure underlying function transformations and by offering tools that support deeper reasoning for learners, this work contributes to a more conceptually coherent and pedagogically effective treatment of a foundational topic in high school mathematics.

Future research may extend this framework in several productive directions. Empirical studies could examine how students' conceptual understanding evolves when instruction emphasizes derivation and reasoning rather than memorization. The same framework may also be adapted for more advanced topics—such as piecewise functions, transformations in the complex plane, matrix-based transformations in linear algebra, or Fourier transformation for digital signal processing—to explore its scalability across mathematical domains. Additionally, partnerships with classroom teachers could inform the development of instructional materials, dynamic software activities, or assessment tasks that further operationalize this approach in diverse learning settings.

■ Acknowledgments

Special appreciation goes to my mentor, who provided extensive support toward the completion of this research project.

■ References

1. Oehrtman, M.; Carlson, M.; Thompson, P. Foundational reasoning abilities that promote coherence in students' function understanding. In *Making the Connection: Research and Teaching in Undergraduate Mathematics Education*; Carlson, M. P.; Rasmussen, C., Eds.; Mathematical Association of America: Washington, DC, **2008**; pp 27–41. <https://pat-thompson.net/PDFversions/2008MAAFunc-tions.pdf>.
2. Daher, W.; Anabousi, A. Students' recognition of function transformations' themes associated with the algebraic representation. *J. Res. Math. Educ.* **2015**, *4* (2), 179–194. <https://doi.org/10.17583/redimat.2015.1110>.

3. Li, X. Mathematical knowledge for teaching algebraic routines: A case study of solving quadratic equations. *J. Math. Educ.* **2011**, 4 (2), 1–16.
4. Telegin, M.; Máté-Klátyik, A. The impact of conceptual versus memorization-based teaching methods on student performance. *J. Emerg. Investig.* **2025**, 1–6. <https://doi.org/10.59720/24-254>.
5. Mocko, M.; Wagler, A.E.; Lesser, L.M.; Francis, W.S.; *et al.* What they remember may not be what they understand: A study of mnemonic recall and performance by introductory statistics students. *J. Stat. Data Sci. Educ.* **2024**, 32 (4), 1–28. <https://doi.org/10.1080/26939169.2024.2334905>.
6. Faulkenberry, E.; Faulkenberry, T. Transforming the way we teach function transformations. *Math. Teach.* **2010**, 104 (1), 29–33. <https://doi.org/10.5951/mt.104.1.0029>.
7. National Council of Teachers of Mathematics. *The Mathematics Teacher* **1980**, 73 (9), 713–714. <https://doi.org/10.5951/mt.73.9.0713>.
8. National Council of Teachers of Mathematics. *Content Standards for School Mathematics*; NCTM: Reston, VA, 2025. <https://staff-old.najah.edu/sites/default/files/content%20standards%20for%20school%20mathematics.pdf>.
9. Borba, M.; Confrey, J. A student's construction of transformations of functions in a multiple representational environment. *Educ. Stud. Math.* **1996**, 31 (3), 319–337. <https://doi.org/10.1007/BF00376325>.
10. Smith, H. *Teaching and Learning Function Transformations*; Arizona State University: Tempe, AZ, 2009. <https://pat-thompson.net/PDFversions/Theses/2009Smith.pdf>.
11. Ng, B.; Hu, C. Use web-based simulation to learn trigonometric curves. *Int. J. Math. Teach. Learn.* **2006**. <https://www.cimt.org.uk/journal/chunhu.pdf>.
12. O'Callaghan, B. Computer-intensive algebra and students' conceptual knowledge of functions. *J. Res. Math. Educ.* **1998**, 29 (1), 21–40. <https://doi.org/10.2307/749716>.
13. Stylianides, A.; Ball, D. Understanding and describing mathematical knowledge for teaching: Knowledge about proof for engaging students in the activity of proving. *J. Math. Teach. Educ.* **2008**, 11 (4), 307–332. <https://doi.org/10.1007/s10857-008-9077-9>.
14. Wu, H. Basic skills versus conceptual understanding. *Am. Educ.* **1999**, 23 (3), 14–19. <https://math.berkeley.edu/~wu/wu1999.pdf>.
15. McCormick, R. Conceptual and procedural knowledge. *Int. J. Technol. Des. Educ.* **1997**, 7 (1), 141–159. <https://doi.org/10.4324/9780203063538>.

■ Author

Elizabeth Kong is a sophomore at Northview High School in Georgia. She has a strong interest in studying mathematical problems and is passionate about identifying effective learning methods to support her peers. In addition to her academic pursuits, Elizabeth is a dedicated violinist, an active debater, and an influential business case study presenter. She aspires to study mathematics, finance, and business leadership at a top university.

Etomidate-Induced Cortical Inhibition: Prefrontal–Amygdala and Its Implications for Emotion and Behavior

Joshua Goh

UWC Southeast Asia (Dover Campus) 1207 Dover Rd, Level 2, 139654, Singapore; khkang74@hotmail.com

Mentor: Xing Yan Choo

ABSTRACT: Etomidate is a widely used anesthetic that produces hypnosis through positive allosteric modulation of γ -aminobutyric acid type A (GABA_A) receptors, enhancing inhibitory neurotransmission and suppressing cortical activity. Growing evidence suggests that etomidate's effects extend beyond sedation, involving disruption of cortical-limbic regulatory circuits. Inhibition within the prefrontal cortex (PFC) may impair top-down modulation of the amygdala, a pathway critical for emotional and behavioral regulation. This narrative literature review synthesizes mechanistic and clinical evidence on how etomidate-induced cortical inhibition affects prefrontal–amygdala circuitry and its implications for cognition, emotion, and behavior. Peer-reviewed experimental and clinical studies examining GABA_A receptor modulation, cortical inhibition, and prefrontal–amygdala connectivity following etomidate exposure were reviewed. The literature indicates that etomidate enhances both synaptic and extrasynaptic GABA_A receptor activity, producing a sustained inhibitory tone that suppresses PFC function and weakens executive control over limbic processing. Excessive inhibition may disrupt prefrontal–amygdala balance, potentially contributing to emotional dysregulation and maladaptive behavioral responses in a dose, duration, and context-dependent manner. This review highlights the importance of considering etomidate circuit-level effects on prefrontal–amygdala interactions when evaluating its broader neuropsychological and neurobehavioral consequences.

KEYWORDS: Medical & Neuroscience, Etomidate & GABAergic Modulation, Amygdala and Prefrontal-limbic Connectivity, Cortical Inhibition.

■ Introduction

Etomidate is a general anesthetic with hypnotic effects mediated through GABA_A receptor activity. It exerts distinct effects on both synaptic and extrasynaptic GABA_A receptors, enhancing inhibitory neurotransmission. Although the precise structural details of its binding site remain limited, growing evidence suggests that etomidate interacts with the interface between the receptor's α and β subunits within the transmembrane domain.¹ This interaction amplifies GABA_A receptor function, resulting in sustained suppression of neuronal excitability and a marked reduction in cortical activity.

Within the mammalian neocortex, billions of neurons form dense networks interconnected through chemical synapses. These networks are organized into functional processing units known as microcircuits and support large-scale parallel information processing.² Each microcircuit consists of tightly coupled excitatory and inhibitory neurons that determine how information is transmitted across cortical regions. Because the prefrontal cortex (PFC) depends on a finely regulated balance of excitation and inhibition to regulate downstream limbic structures, understanding how etomidate enhances GABA_A-mediated inhibition is essential for explaining its effects on cognition, emotion, and behavior.

Functional specialization within the PFC further underscores its central role in higher-order regulation. Distinct prefrontal subregions contribute differentially in cognitive control, emotional appraisal, and behavioral inhibition. The dorsolateral PFC is primarily associated with executive functions such as

working memory and cognitive reappraisal, whereas the ventromedial PFC contributes to emotional valuation and the inhibition of fear-related responses. Through these functions, the PFC exerts top-down modulatory influence over subcortical limbic structures, most notably the amygdala. Disruption of prefrontal–amygdala connectivity is implicated in affective and behavioral disturbances, including anxiety and mood disorders.^{3–6}

The amygdala plays a central role in detecting emotionally salient stimuli and generating affective responses, particularly those related to fear, threat, and stress.⁴ Effective emotional and behavioral regulation depends on intact prefrontal–amygdala communication, whereby PFC activity constrains or modulates amygdala reactivity in a context-dependent manner. Disruption of this regulatory circuitry has been linked to maladaptive emotional responses and behavioral dysregulation.

Given etomidate's potent enhancement of GABA_A-mediated inhibition, its effects on PFC microcircuits are likely to have downstream consequences for prefrontal–amygdala interactions. By increasing inhibitory tone within the PFC, etomidate may alter the balance between cortical control and limbic reactivity, thereby influencing emotional processing and behavioral output.⁷ These effects suggest that the clinical and non-clinical use of etomidate and its neuropsychological impacts extend beyond generalized hypnosis and may involve selective modulation of regulatory brain networks.⁸ This review synthesizes evidence from both clinical administration and emerging non-medical misuse of etomidate to examine how

etomidate-induced cortical inhibition modulates prefrontal–amygdala circuitry and influences cognitive, emotional, and behavioral regulation.

■ Methodology

Studies relevant to etomidate-induced cortical inhibition and its effects on prefrontal–amygdala circuitry and emotional regulation were included in this literature review. Literature searches were conducted using PubMed and Google Scholar between January 2025 and December 2025.

Search terms included combinations of “etomidate”, “GABA_A receptor”, “cortical inhibition”, “synaptic inhibition”, “extrasynaptic GABA receptors”, “prefrontal cortex”, “amygdala”, “prefrontal–amygdala connectivity”, and “emotional regulation”. Boolean operators (AND/OR) were used to refine searches and capture mechanistic, preclinical, and clinical studies relevant to the aims of the review. Studies were selected based on relevance to the review topic, regardless of publication year.

Inclusion criteria:

Studies were included if they were peer-reviewed, available in full text, and provided mechanistic, experimental, or observational data on etomidate, GABA_A receptor modulation, or cortical inhibition. Articles were also required to examine the prefrontal cortex, amygdala, or their connectivity, and to report findings relevant to emotional or behavioral regulation. Only studies published in English were considered.

Exclusion criteria:

Studies were excluded if they were not peer-reviewed, not available in full text, or focused solely on peripheral or non-cortical effects of etomidate, such as adrenal suppression. Articles that did not examine GABA_A receptor modulation, cortical inhibition, or brain-based mechanisms relevant to emotional regulation were also excluded. Preprints, conference abstracts, and studies unrelated to the prefrontal cortex, amygdala, or associated neural circuitry were omitted.

■ Results and Discussion

This review aims to underscore the significance of etomidate circuit-level effects on prefrontal–amygdala interactions when assessing the neurological impacts on humans. By understanding these mechanisms, etomidate-related issues are urgent and relevant, highlighting why it deserves immediate attention and informed action in light of emerging reports of non-medical etomidate exposure. This highlights the need for further experimental research on its long-term cognitive, emotional, and behavioral consequences.

Etomidate Modulation of GABA_A Receptors:

Etomidate exerts its primary neuropharmacological effects through potentiation of γ -aminobutyric acid type A (GABA_A) receptors, resulting in enhanced inhibitory neurotransmission across cortical and subcortical networks, as shown in Figure 1. Experimental studies using GABAergic inhibitory models

support this mechanism. Muscimol, a direct GABA_A receptor agonist, administered at concentrations ranging from 0.1 μ M to 100 μ M, increases GABA_A-mediated chloride ion influx, leading to neuronal hyperpolarization, reduced neuronal excitability, and widespread suppression of neural activity by using an *in vitro* neurons model to examine chloride flux across isolated brain membranes following muscimol exposure, functionally similar to etomidate.⁹ These findings support the mechanistic role of enhanced GABAergic chloride conductance in mediating neuronal inhibition, although the experimental membrane preparation may not fully replicate physiological or clinical etomidate exposure.

Similarly, an *in vivo* electrophysiological study using intact feline models demonstrated that propofol and etomidate, used without reported doses or concentrations, significantly reduced neuronal firing activity across cortical, thalamic, and reticular regions during intravenous anesthetic exposure, accompanied by electroencephalography (EEG) changes consistent with anesthetic-induced unconsciousness.¹⁰ These findings support the concept that widespread GABAergic inhibition disrupts higher-order neural integration and cortical information processing. However, the relatively deep anesthetic exposure used in the study may exceed typical clinical or recreational etomidate exposure.

At the cellular level, etomidate enhances both synaptic and extrasynaptic GABA_A receptor activity, producing sustained inhibitory tone beyond normal phasic GABA signaling. At higher concentrations, etomidate can directly activate GABA_A receptor-associated chloride channels even in the absence of endogenous GABA, as shown in Figure 1, resulting in profound neuronal hyperpolarization and severe suppression of cortical, thalamic, and limbic activities.^{9,10} These findings provide a mechanistic basis for the downstream effects of etomidate on higher-order neural regulation.

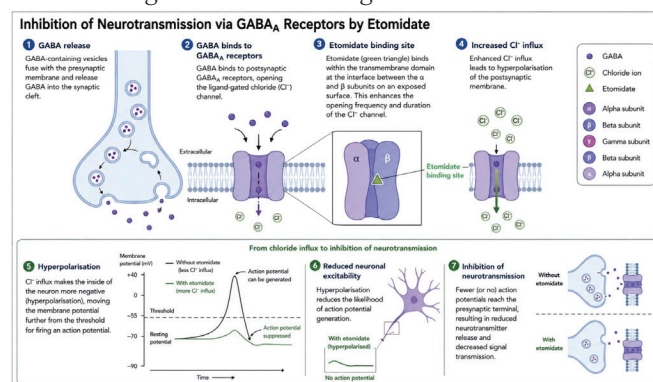


Figure 1: Etomidate enhances GABA_A receptor-mediated inhibition. (1) GABA is released from the presynaptic neuron. (2) GABA binds to postsynaptic GABA_A receptors, opening ligand-gated chloride (Cl⁻) channels. (3) Etomidate (green triangle) binds at the interface between the α and β subunits within the transmembrane domain, enhancing channel opening. (4) This increases Cl⁻ influx. (5) The resulting Cl⁻ influx hyperpolarises the postsynaptic membrane. (6) Hyperpolarisation reduces neuronal excitability and the likelihood of action potential generation. (7) Consequently, less neurotransmitter is released, leading to inhibition of neurotransmission. Created by the author, Joshua Goh, under the guidance of his mentor, Xing Yan Choo.

Etomidate Effects on Prefrontal Cortex:

The prefrontal cortex (PFC) plays a central role in emotional regulation by exerting top-down control over subcortical structures, most notably the amygdala. This regulatory function supports cognitive reappraisal, impulse control, and adaptive decision-making, depending on a finely tuned balance of excitation and inhibition within prefrontal microcircuits.¹¹⁻¹³ Etomidate-induced enhancement of GABA_A-mediated inhibition has the potential to disrupt this balance by suppressing neuronal firing and weakening functional connectivity within prefrontal networks.

Distinct PFC subregions contribute differentially to emotional regulation. The dorsolateral PFC supports executive control and cognitive reappraisal, whereas ventromedial and ventrolateral regions are involved in emotional valuation and inhibition of fear-related responses.¹³ Suppression of activity across these regions may weaken executive oversight of emotional processing, increasing vulnerability to maladaptive emotional responses, impaired decision-making, and reduced behavioral inhibition.

Neurophysiological evidence suggests that etomidate can impair synaptic plasticity and reduce frontal network coherence, both of which are essential for dynamic prefrontal regulation.^{14,15} In addition to these direct cortical effects, etomidate may indirectly compromise PFC function through suppression of cortisol production. Even a single dose has been shown to reduce cortisol levels for up to 24 hours,¹⁶ potentially impairing stress adaptation and further weakening prefrontal-mediated inhibition of limbic activity.

Overall, etomidate-induced suppression of prefrontal microcircuits may compromise executive control and emotional regulation under conditions requiring adaptive cognitive or stress-related responses. Limbic dysfunction can not only manifest as emotional and behavioral change, but also affect an individual with various psychological conditions.

Prefrontal-Amygdala Interactions under Etomidate:

The prefrontal-amygdala circuit represents a core neural substrate for emotional regulation and behavioral control. Within this circuitry, the PFC implements cognitive strategies that modulate emotional responses, while the amygdala processes emotional salience, fear, and affective memory.¹⁷ The PFC, connected to the limbic system, regulates emotional responses by exerting inhibitory control over the system, a network of subcortical structures, through inhibitory modulation of limbic structures, particularly the amygdala.¹⁸ Effective emotional regulation depends on intact bidirectional communication between these regions, allowing contextual information processed by the PFC to constrain amygdala reactivity.¹⁹

Etomidate-induced cortical inhibition may disrupt this regulatory balance by preferentially suppressing prefrontal activity relative to limbic responsiveness. Excessive inhibition within the PFC can weaken top-down control, allowing amygdala-driven emotional and behavioral responses to dominate, as shown in Figure 2.

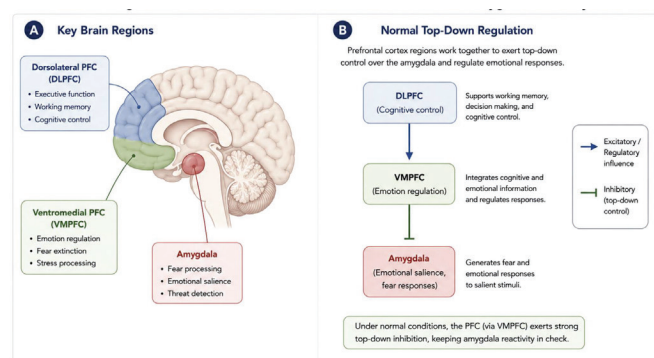


Figure 2: (A) Key brain regions involved in prefrontal-amygdala circuitry. The dorsolateral prefrontal cortex (DLPFC) supports cognitive control and executive functions, while the ventromedial prefrontal cortex (VMPFC) regulates emotions and stress responses. The amygdala mediates fear and emotional salience. (B) Under normal conditions, the DLPFC supports cognitive control through engagement of the VMPFC, which modulates and suppresses amygdala activity. This top-down regulation prevents excessive emotional reactivity and maintains adaptive behavioral responses. Created by the author, Joshua Goh, under the guidance of his mentor, Xing Yan Choo.

Consistent with this mechanism, experimental studies in animal models have shown that prolonged or excessive etomidate exposure disrupts neurotransmitter signaling, compromises neural integrity, and induces behavioral alterations, including heightened aggression, psychiatric-like manifestations, and neurotoxicity, with evidence of irreversible nervous system damage and increased mortality in severe cases.²⁰ Importantly, the effects of etomidate on prefrontal-amygdala connectivity appear to be highly dose-, duration-, and context-dependent. Acute exposure in infant mice did not result in long-term behavioral or motor impairments and was associated with recovery-related hippocampal changes.²¹ Clinical studies similarly suggest that carefully titrated etomidate use may preserve cognitive function, with reduced postoperative cognitive dysfunction observed in elderly patients receiving etomidate alone or in combination with propofol.²²

Etomidate may also indirectly disrupt prefrontal-amygdala regulation through cortisol suppression, as reduced cortisol weakens medial PFC inhibition of the amygdala and impairs stress regulation.²³ As a result, disruption of prefrontal-amygdala balance represents a plausible circuit-level mechanism by which excessive or prolonged etomidate exposure may contribute to emotional dysregulation and maladaptive behavioral outcomes. In addition, etomidate is also frequently associated with extrapyramidal-like motor manifestations, particularly myoclonus. Although these movements may resemble extrapyramidal symptoms clinically, current evidence suggests they arise from transient cortical disinhibition and altered glutamatergic activity rather than classical basal ganglia dopaminergic dysfunction.²⁴⁻²⁶ Etomidate induces myoclonus through rapid cortical inhibition preceding subcortical suppression, resulting in disinhibition of motor pathways.²⁷ As a GABAergic anesthetic, etomidate also suppresses arousal networks and can induce loss of consciousness through enhanced inhibitory neurotransmission.²⁵ Acute inhalation of etomidate-laced e-cigarettes, an emerging concern, leads to CNS depression and adrenal insufficiency by inhibiting adrenal cortical

enzymes, resulting in unresponsiveness and loss of consciousness.²⁶

Implications for Emotional and Behavioral Regulation:

Collectively, these findings indicate that etomidate influences emotional and behavioral regulation through converging effects on GABAergic inhibition, PFC function, cortical–limbic connectivity, and neuroendocrine modulation. By enhancing inhibitory signaling, etomidate suppresses PFC activity and weakens top-down regulatory control over the amygdala. This shift in cortical–limbic balance increases vulnerability to emotional dysregulation, heightened stress reactivity, and impaired emotional reappraisal, particularly in emotionally salient or stressful contexts.

From a behavioral perspective, disruption of PFC-mediated executive control may result in increased impulsivity, maladaptive decision-making, and difficulty regulating goal-directed behavior. Experimental models demonstrating aggression, behavioral disturbances, and neurotoxicity following chronic and high-dose etomidate exposure support the link between impaired cortical regulation and adverse behavioral outcomes.

These effects may be further amplified by etomidate-induced suppression of cortisol, a hormone essential for stress adaptation and emotional stability.²⁸ Reduced neuroendocrine support of prefrontal–limbic circuitry may therefore exacerbate deficits in emotional and behavioral regulation. Importantly, the neurobehavioral effects of etomidate are highly context-dependent. Acute or carefully titrated exposure in clinical settings may preserve, or even facilitate, cognitive function and recovery. In contrast, chronic exposure or supratherapeutic dosing may compromise neural integrity and behavioral control.

Excessive etomidate exposure can produce exaggerated GABAergic inhibition, resulting in profound sedative effects and cumulative neuroendocrine disruption, particularly when used outside controlled clinical supervision.²⁹ Collectively, these observations highlight the importance of considering both neural circuit dynamics and neuroendocrine context when evaluating etomidate’s effects on emotional and behavioral regulation.

Research in some animals and humans demonstrates comparable prefrontal executive functions that regulate cognitive processes critical for behavioral inhibition, planning, decision-making, etc.³⁰ Experimental animal studies, summarized in Table 1, further indicate an increased risk of adverse neurological and behavioral outcomes in humans under such conditions, as summarized in Table 2. It is noted that the homology of prefrontal executive functions across animals and humans facilitates numerous translational studies using animal models of emotion and behavior regulation, as displayed in Table 1. For example, the emerging finding on the extinction of conditioned fear is elucidated, regardless of the challenge of researching the regulation of emotion in rodents.³¹

Table 1A: Animal evidence from behavioral and emotional domains relevant to amygdala-related GABAergic function. These animal studies exhibit key emotional domains, including fear regulation, stress, and behavioral issues such as aggression, unconsciousness, and developmental vulnerability, holistically demonstrating that GABAergic modulation strongly influences amygdala-related emotional and behavioral malfunctions.

Behavioral/Emotional Domain	References	Species/Model	Brain Region/System	GABAergic Mechanism	Key Behavioral/Emotional Findings	Translational Relevance
Fear, Anxiety, and Emotion Regulation	Quirk & Beer (2006) 17	Rat	Prefrontal–amygdala circuit	GABAergic inhibitory regulation of amygdala	Enhanced prefrontal inhibition reduces fear responses	Converges with human emotion regulation models
Anxiety	Allan & Harris (1986) 9	Rat brain membranes	Brain membranes	GABA antagonist effects on Cl ⁻ flux	Altered inhibition affects neural excitability linked to anxiety	Mechanistic support for anxiety modulation
Emotion	Kohl & Paulsen (2010) 2	Rodent	Cortical networks	GABA _A receptor modulation	Stabilization of neural network activity in emotional control	Supports connectivity findings in humans
Stress, Arousal, and Network Stability	Talavera <i>et al.</i> (2009) 7	Rodent	Thalamocortical system	Etomidate potentiation of GABA _A receptors	Reduced arousal and disrupted cortical integration	Exhibits stress-related neural dampening
Stress	Andrada <i>et al.</i> (2012) 11	Rat	Cortex, reticular formation	GABAergic anesthetic suppression	Decreased neuronal firing and responsiveness	Highlights Relevant to stress and affective blunting
Aggression and Affective Dysregulation	Li <i>et al.</i> (2025) 14	Zebrafish	Whole brain	Etomidate-induced GABAergic neurotoxicity	Increased aggression and behavioral dysregulation	Parallels emotional regulation in humans
Developmental Emotional Outcomes	Nyman <i>et al.</i> (2016) 21	Infant mice	Central nervous system	GABAergic anesthetic action and inhibitory CNS	No significant behavioral or emotional impairment	Elucidates developmental resilience
Dysregulation	Ding <i>et al.</i> (2024) 20	Mouse	Nervous system gut-brain axis	Chronic GABAergic disruption	Long-term neural and systemic dysregulation	Warns Prolonged exposure effects on the nervous system
Consciousness and Emotional Blunting	Talavera <i>et al.</i> (2009) 7	Rodent	Thalamocortical networks	GABA _A mediated suppression	Reduced conscious and emotional responsiveness	Relevant to anesthetic-induced injurious effects on consciousness
Emotion	Andrada <i>et al.</i> (2012) 11	Rat	Limbic–cortical networks	Global GABAergic depression	Emotional regulation under anesthesia	Explains reduced affect in humans

The findings presented in Table 1 are refined as follows: GABA_A receptor-mediated inhibition in PFC induced by low-to-moderate etomidate does not directly increase emotional reactivity, such as fear, anxiety, or impulsive responses. However, when excessive and prolonged etomidate is administered, limbic circuits become disinhibited. Decreased GABAergic inhibition makes Cl⁻ influx less, rendering neurons more excitable. As a result, increased anxiety-like behaviors, such as impulsiveness or aggression, are found.

Table 1B: Summary of animal studies examining neural mechanisms, levels of dysfunction, and reported neuronal damage or cell death associated with etomidate exposure. Findings range from minimal functional impairment under low-to-moderate exposure to neurotoxicity, neural disruption, and behavioral dysregulation following excessive or prolonged exposure, highlighting the dose- and duration-dependent nature of etomidate’s effects.

Behavioral/Emotional Domain	References	Exposure Level	Neural Mechanism	Level of Dysfunction	Neuronal Damage / Cell Death	Reported Key Findings
Fear, Anxiety, and Emotion Regulation	Quirk & Beer (2006) 17	Not reported	Enhanced prefrontal inhibitory regulation	Circuit-level	Not reported	Enhanced prefrontal inhibition reduced fear responses
Anxiety	Allan & Harris (1986) 9	Not reported	Altered inhibitory neural excitability	Synaptic/ Circuit-level	Neuronal damage	Altered inhibition affected neural excitability linked to anxiety

Emotion	Kohli & Paulsen (2010) 2	Not reported	Stabilization of neural network activity	Network-level	Not reported	Stabilization of neural network activity involved in emotional control
Developmental Emotional Outcomes	Nyman <i>et al.</i> (2016) 21	Unclearly reported (clinically induced unconsciousness level)	Minimal developmental neural disruption	Systems-level	No significant neuronal damage reported	No significant behavioral or emotional impairment observed
Behavioral/Emotional Dysregulation	Ding <i>et al.</i> (2024) 20	Excessive (chronic exposure)	Chronic neural disruption and hyperexcitability	Systems-level	Neuronal damage	Increased etomidate-induced myoclonus
Stress, Arousal, and Network Stability	Talavera <i>et al.</i> (2009) 7	Unclearly reported (clinically relevant anesthetic exposure)	Disrupted cortical integration and suppression	Network-level	Functional suppression reported; cell death not specified	Reduced arousal and disrupted cortical integration
Stress	Andrada <i>et al.</i> (2012) 11	Moderate / Excessive	Suppressed neuronal firing responsiveness	Circuit-level	Neuronal suppression reported; cell death not specified	Decreased neuronal firing and responsiveness
Aggression and Affective Dysregulation	Li <i>et al.</i> (2025) 14	Not reported	Neurotoxicity and apoptosis-associated dysregulation	Molecular/ Systems-level	Neuronal damage; neurotoxicity and apoptosis reported	Increased aggression and behavioral dysregulation
Dysregulation	Ding <i>et al.</i> (2024) 20	Excessive	Chronic systemic neural disruption	Systems-level	Chronic neural disruption reported	Long-term neural and systemic dysregulation
Consciousness and Emotional Blunting	Talavera <i>et al.</i> (2009) 7	Not reported	Functional cortical suppression	Network-level	Functional cortical suppression reported	Reduced consciousness and emotional responsiveness
Emotion	Andrada <i>et al.</i> (2012) 11	Unclearly reported (clinically relevant anesthetic exposure)	Global neural suppression under anesthesia	Systems-level	Global neural suppression reported	Emotional regulation suppression under anesthesia

However, human research conducted to verify etomidate's potential effects is constrained by underlying risks. Only a few studies have been conducted, leading to the limited findings, thereby providing inconclusive findings regarding emotional, cognitive, and behavioral effects, as reflected in Table 2.

Table 2: Statistically reported cognitive, emotional, and affective outcomes following etomidate administration in human subjects, with respective study models.

Domain	References	Study Model	Key Findings	Reported Statistics
Cognitive function	Zhi & Li (2023) 22; Absalom <i>et al.</i> (1999) 16	Older adults	Transient postoperative cognitive dysfunction	Incidence happens in 25–30% within 7 days post-anesthesia
Emotional state	Zhi & Li (2023) 22	Human through experiments	Increased anxiety and depressive symptoms following anesthesia	Significant increases vs. baseline ($p < 0.05$)
Emotional regulation	Veer <i>et al.</i> (2012) 23	Human through neuroimaging	Altered prefrontal–amygdala functional connectivity	Significant connectivity changes associated with cortisol levels ($p < 0.05$)
Affective processing	Banks <i>et al.</i> (2007) 12; Wager <i>et al.</i> (2008) 13	Humans through functional imaging	Disrupted prefrontal–limbic interactions relevant to emotion control	Significant network-level modulation vs. baseline

■ Limitation

First, for ethical reasons, direct investigation of etomidate's effects on cognition, emotional regulation, and behavior in humans is inherently limited. Deliberate exposure of participants to potentially harmful doses or prolonged use of etomidate is not permissible, restricting human studies largely to clinical, observational, or case-based evidence.

Second, data regarding non-medical and recreational etomidate use remain limited. Most available evidence originates from case reports, toxicology studies, and observational investigations, making it difficult to fully characterize patterns of misuse, long-term neurobehavioral effects, and associated risks. Furthermore, unlike clinical anesthetic use, in which etomidate is administered intravenously at standardized and well-characterized doses, the actual dose exposure associated with inhalational use remains poorly defined. Variability in product formulations, inhalation patterns, and frequency of use, along with the lack of standardized toxicological analyses, makes exposure estimation difficult. Similarly, epidemiological data regarding duration of use and peak-age distribution among users remain limited, restricting conclusions about dose-response relationships and long-term risk.

Third, the neurobehavioral risks associated with etomidate are likely to vary across patient populations. Vulnerable groups, including elderly individuals and patients with chronic medical conditions, may be at increased risk of adverse cognitive, emotional, or physiological effects, limiting the generalizability of findings derived from healthier populations.

Finally, severe neuropsychiatric outcomes such as mental disorders, self-harm behaviors, and fatal events have been reported in cases of extreme etomidate exposure. However, systematic research on etomidate toxicity remains limited due to the scarcity of controlled studies, inconsistent reporting of poisoning manifestations, and the lack of standardized assessment protocols. These gaps highlight the need for more structured toxicological and clinical investigations.

■ Recommendations and Future Directions

Recent reports indicate that etomidate has emerged as a misused psychoactive substance, drawing increasing attention due to its presence in certain illicit formulations, including products such as K-POD. This trend has heightened concern regarding non-medical exposure to etomidate and reinforced the urgency of further investigation to understand its neurobiological, emotional, and behavioral consequences. The potential for abuse-related outcomes, including irreversible neural damage and emotional or behavioral disturbances, underscores the need for targeted preventive and public health measures.

Future research should prioritize elucidating the mechanisms underlying etomidate-related morbidity and mortality, particularly in cases of suspected poisoning. While adrenocortical suppression and myoclonus are recognized contributors to adverse outcomes, additional mechanisms may be involved. Integrated neuroendocrine, neuroimaging, and clinical studies are needed to directly assess amygdala–prefrontal cortex dysfunction, cortisol suppression, and their roles in cognitive, emotional, and behavioral impairments associated with etomidate toxicity.

■ Conclusion

Longitudinal and translational studies examining repeated or chronic etomidate exposure are needed to more clearly char-

acterize its long-term neuropsychological effects, particularly in contexts extending beyond controlled clinical use. Disruption of emotional regulation and behavioral control may carry meaningful psychosocial consequences, including impaired interpersonal functioning and increased risk of maladaptive behaviors. A clearer understanding of these risks would support improved clinical risk stratification, inform regulatory and preventive strategies, and provide a foundation for targeted interventions aimed at minimizing potential harm associated with etomidate-containing substances. Greater awareness of the neural, behavioral, and neuroendocrine consequences of etomidate exposure is therefore essential to ensure its safe and responsible use.

■ Acknowledgments

The author would like to thank the mentors and educators who guided the development of this narrative literature review. Special appreciation is extended to those who offered feedback on the research structure, scientific writing, and interpretation of findings. The author is also grateful to the researchers whose published work formed the foundation of this review. Finally, sincere thanks are given to family and peers for their encouragement and support throughout the research and writing process.

■ References

- Chiara, D. C.; Dostalova, Z.; Jayakar, S. S.; Zhou, X.; Miller, K. W.; Cohen, J. B. Mapping general anesthetic binding site(s) in human $\alpha 1\beta 3$ γ -aminobutyric acid type A receptors with [(3)H]TDB-zl-etomidate, a photoreactive etomidate analogue. *Biochemistry* **2012**, *51*, 836–847. <https://doi.org/10.1021/bi201772m>
- Kohl, M. M.; Paulsen, O. The roles of GABA_A receptors in cortical network activity. *Adv. Pharmacol.* **2010**, *58*, 205–229. [https://doi.org/10.1016/S1054-3589\(10\)58009-8](https://doi.org/10.1016/S1054-3589(10)58009-8)
- Hiser, J.; Koenigs, M. The multifaceted role of the ventromedial prefrontal cortex in emotion, decision making, social cognition, and psychopathology. *Biol. Psychiatry* **2018**, *83* (8), 638–647. <https://doi.org/10.1016/j.biopsych.2017.10.030>
- Berboth, S.; Morawetz, C. Amygdala-prefrontal connectivity during emotion regulation: A meta-analysis of psychophysiological interactions. *Neuropsychologia* **2021**, *153*, 107767. <https://doi.org/10.1016/j.neuropsychologia.2021.107767>
- Kovner, R.; Oler, J. A.; Kalin, N. H. Cortico-limbic interactions mediate adaptive and maladaptive responses relevant to psychopathology. *Am. J. Psychiatry* **2019**, *176* (12), 987–999. <https://doi.org/10.1176/appi.ajp.2019.19101064>
- Hosseini, S.; Cooper, J. A.; DeVries, B. A.; Nuutinen, M. R.; Hahn, E. C.; Kragel, P. A.; Treadway, M. T. Effects of acute stress and depression on functional connectivity between prefrontal cortex and the amygdala. *Mol. Psychiatry* **2023**, *28* (11), 4602–4612. <https://doi.org/10.1038/s41380-023-02056-5>
- Talavera, J. A.; Esser, S. K.; Amzica, F.; Hill, S.; Antognini, J. F. Modeling the GABAergic action of etomidate on the thalamo-cortical system. *Anesth. Analg.* **2009**, *108* (1), 160–167. <https://doi.org/10.1213/ane.0b013e31818d40aa>
- Uhm, J.; Hong, S.; Han, E. The need to monitor emerging issues in etomidate usage: the misuse or abuse potential. *Forensic Sci. Med. Pathol.* **2024**, *20* (1), 249–260. <https://doi.org/10.1007/s12024-023-00596-4>
- Allan, A. M.; Harris, R. A. γ -Aminobutyric acid agonists and antagonists alter chloride flux across brain membranes. *Mol. Pharmacol.* **1986**, *29* (5), 497–505. [https://doi.org/10.1016/S0026-895X\(25\)10278-2](https://doi.org/10.1016/S0026-895X(25)10278-2)
- Wang, Z.; Li, A.; Xiao, G.; Gu, X.; et al. Unbiased segmentation and multifeature classification of cortical neuronal activity reveal complex dynamics under anesthesia. *Neurosci. Bull.* **2025**, <https://doi.org/10.1007/s12264-025-01527-9>
- Andrada, J.; Livingston, P.; Lee, B. J.; Antognini, J. Propofol and etomidate depress cortical, thalamic, and reticular formation neurons during anesthetic-induced unconsciousness. *Anesth. Analg.* **2012**, *114* (3), 661–669. <https://doi.org/10.1213/ANE.0b013e3182405228>
- Banks, S. J.; Eddy, K. T.; Angstadt, M.; Nathan, P. J.; Phan, K. L. Amygdala-frontal connectivity during emotion regulation. *Soc. Cogn. Affect. Neurosci.* **2007**, *2* (4), 303–312. <https://doi.org/10.1093/scan/nsm029>
- Wager, T. D.; Davidson, M. L.; Hughes, B. L.; Lindquist, M. A.; Ochsner, K. N. Prefrontal-subcortical pathways mediating successful emotion regulation. *Neuron* **2008**, *59* (6), 1037–1050. <https://doi.org/10.1016/j.neuron.2008.09.006>
- Li, X.; Liu, J.; Sun, Y.; Wang, H.; Zhang, Q. Neurotoxicity and aggressive behavior induced by anesthetic etomidate exposure in zebrafish: Insights from multi-omics and machine learning. *Aquat. Toxicol.* **2025**, *282*, 107321. <https://doi.org/10.1016/j.aquatox.2025.107321>
- Vesuna, S.; Kauvar, I. V.; Richman, E.; Gore, F.; Oskotsky, T.; Sava-Segal, C.; ... Deisseroth, K. Deep posteromedial cortex rhythm in dissociation and depression. *Nature* **2020**, *586* (7827), 87–94. <https://doi.org/10.1038/s41586-020-2731-9>
- Absalom, A.; Pledger, D.; Kong, A. Adrenocortical function in critically ill patients 24 h after a single dose of etomidate. *Anaesthesia* **1999**, *54* (9), 861–867. <https://doi.org/10.1046/j.1365-2044.1999.01003.x>
- Quirk, G. J.; Beer, J. S. Prefrontal involvement in the regulation of emotion: convergence of rat and human studies. *Curr. Opin. Neurobiol.* **2006**, *16*, 723–727. <https://doi.org/10.1016/j.conb.2006.07.004>
- Etkin, A.; Wager, T. D. Functional Neuroimaging of Anxiety: A Meta-Analysis of Emotional Processing in PTSD, Social Anxiety Disorder, and Specific Phobia. *Am. J. Psychiatry* **2007**, *164* (10), 1476–1488. <https://doi.org/10.1176/appi.ajp.2007.07030504>
- Ochsner, K. N.; Ray, R. D.; Cooper, J. C.; Robertson, E. R.; Chopra, S.; Gabrieli, J. D. E.; Gross, J. J. For better or for worse: Neural systems supporting the cognitive down- and up-regulation of negative emotion. *NeuroImage* **2004**, *23* (2), 483–499. <https://doi.org/10.1016/j.neuroimage.2004.06.030>
- Ding, S.; Li, K.; Han, X.; Lin, W.; Qin, Y.; Cao, R.; Ren, Y. Long-term use of etomidate disrupts intestinal homeostasis and the nervous system in mice. *Toxicol.* **2024**, *504*, 153718.
- Nyman, Y.; Fredriksson, A.; Lönnqvist, P. A.; Viberg, H. Etomidate exposure in early infant mice (P10) does not induce apoptosis or affect behaviour. *Acta Anaesthesiol. Scand.* **2016**, *60* (5), 588–596. <https://doi.org/10.1111/aas.12685>
- Zhi, Y.; Li, W. Effects of total intravenous anesthesia with etomidate and propofol on postoperative cognitive dysfunction. *Physiol. Res.* **2023**, *72* (2), 251–258. <https://doi.org/10.3354/physiolres.934983>
- Veer, I. M.; Oei, N. Y. L.; Spinhoven, P.; van Buchem, M. A.; Elzinga, B. M.; Rombouts, S. A. R. B. Endogenous cortisol is associated with functional connectivity between the amygdala and medial prefrontal cortex. *Psychoneuroendocrinology* **2012**, *37* (7), 1039–1047. <https://doi.org/10.1016/j.psyneuen.2011.12.001>

24. Urman, R. D.; Malapero, R. J.; Zaccagnino, M. P.; Brovman, E. Y.; Kaye, A. D. *Etomidate Derivatives: Novel Pharmaceutical Agents in Anesthesia. J. Anaesthesiol. Clin. Pharmacol.* **2017**, *33* (4), 429–431. DOI: 10.4103/0970-9185.222521.
25. Ries, C. R.; Puil, E. Mechanism of Anesthesia Revealed by Shunting Actions of Isoflurane on Thalamocortical Neurons. *J. Neurophysiol.* **1999**, *81* (4), 1795–1801. <https://doi.org/10.1152/jn.1999.81.4.1795>.
26. Qin, Y.; Lin, H.; Lv, W.; Hong, S.; Huang, Z. Adrenal Insufficiency Associated with Long-Term Use of Electronic Cigarettes Reportedly Containing Etomidate in Two Patients. *Clin. Toxicol.* **2024**, *62* (10), 672–673. <https://doi.org/10.1080/15563650.2024.2393221>.
27. Wang, T.; *et al.* Severe Myoclonus Masking Thyroid Storm Following Etomidate Administration: A Case Report. *Medicine* **2025**, *104* (39), e4298.
28. Zheng, G.; Chen, Y.; Wu, G.; Song, T.; Zou, X.; Nie, Q.; Zhang, P. Review of the hazards and contraindications of etomidate. *Int. J. Toxicol.* **2025**, *44* (3), 245–253. <https://doi.org/10.1177/10915818241297073>
29. Uhm, J.; Hong, S.; Han, E. The need to monitor emerging issues in etomidate usage: the misuse or abuse potential. *Forensic Sci. Med. Pathol.* **2024**, *20* (1), 249–260. <https://doi.org/10.1007/s12024-023-00596-4>
30. Dalley, J. W.; Cardinal, R. N.; Robbins, T. W. Prefrontal executive and cognitive functions in rodents: Neural and neurochemical substrates. *Neuroscience & Biobehavioral Reviews* **2004**, *28* (7), 771–784. <https://www.sciencedirect.com/science/article/abs/pii/S014976340400096X?via%3Dihub>
31. Quirk, G. J.; Beer, J. S. Prefrontal involvement in the regulation of emotion: Convergence of rat and human studies. *Current Opinion in Neurobiology* **2006**, *16* (6), 723–727. <https://doi.org/10.1016/j.conb.2006.07.004>

■ Author

Joshua Goh is in 11th grade at UWC Southeast Asia and President of the Global True Friend Club, an initiative providing support for at-risk individuals through rehabilitation programs, officially recognized by True Friend Foundation <https://truefriendfoundation.com>. He pursues his passion for science and medicine to address global health challenges.

Emerging Downstream Plastic Waste Disposal Solutions in the United States: A Review of Plastic-Modified Pavement

Jasmine S. Liu

Albany High School, 603 Key Route Blvd, Albany, CA, 94706, USA; jasmineliu1608@gmail.com

ABSTRACT: Plastic pollution has been a persistent and urgent global crisis that has surpassed the capacities of current waste management systems. With only about 5-6% of global plastics currently being recycled and microplastic identification in human and environmental samples, there is an urgent need for alternative sustainable waste management systems. A recent downstream recycling technology that has gained recognition is waste plastic-modified roads. This review aims to provide a comprehensive evaluation of plastic roads as a transitioning alternative from downstream waste mismanagement to a circular plastic economy. Factors, including mechanical performance, various modifier contents, potential environmental impacts, economic feasibility, and sustainable ethics, are examined and contextualized within U.S. environmental regulations. Main findings revealed that certain material compositions exerted beneficial effects on the pavement's durability and rheology; however, severe variability arose in storage stability resulting from poor adhesion between polymer and asphalt components. Longitudinal environmental impacts were uncertain, primarily due to unclear long-term circularity of plastic roads, inconsistent findings across toxicity studies, and the absence of environmental benchmarks. Future areas of research should focus on the mitigation of storage stability issues, chemical toxicity content, reclaimed asphalt pavement, and the establishment of MP release thresholds and relevant environmental standards.

KEYWORDS: Materials Science, Polymers, Waste Plastic Modified Asphalt, Chemical Leaching, Economic Feasibility.

■ Introduction

The urbanizing and convenience-driven consumer market has necessitated the rapid expansion of the U.S. plastics industry. Since its invention in 1907 as Bakelite, plastic has revolutionized virtually all industries, including medicine, aerospace, automotive, construction, electronics, and beyond.¹ As a synthetic polymer, plastic offers a unique combination of properties such as lightweight, durability, flexibility, and cost-effectiveness for production. The versatile and economical nature has led to plastic becoming the second most produced material aside from paper. In parallel, this proliferation of the plastic economy has escalated adverse environmental destruction into a staggering global issue.²

Due to a lack of biodegradability, the safe disposal of plastics remains a challenge. Currently, only around 5-6% are successfully recycled.^{2,3} Plastics that are too damaged or have incompatible melting temperatures are typically managed through end-of-pipe methods. According to the National Renewable Energy Laboratory, approximately 85% are landfilled, and 10% are incinerated, which doesn't even account for the 22% of global plastics that are mismanaged.^{4,5} Plastic pollution poses an immense threat to public and environmental health. Over time, plastics break down into smaller particles, reaching microscopic sizes, known as micro- and nano-plastics.⁶ Unlike virgin polymers, plastics contain elevated amounts of chemical additives that are harmful to ecosystems, especially in microplastic form. These chemicals can leach into soil and bodies of water, accelerating ocean acidification, contaminating nutrient cycles, and threatening biodiversity.⁷ Studies show that microplastics have already been found in the soil, air, oceans, and

human placentas.⁸ Microplastics pose extreme toxicity risks when present in the human body, causing increased hormone fluctuation and risk for neurodegenerative diseases and certain cancers.^{9,10}

With plastic waste projected to triple current production rates and accumulate 140 million metric tons by 2060, there is a rapidly growing disparity between production and disposal.¹¹ U.S. landfills are increasingly burdened by non-recyclable and single-use plastics, contributing to capacity strains. In 2023, the Illinois Environmental Protection Agency forecasted that state-wide landfill capacity would be reached in approximately 19 years.¹² End-of-pipe landfills are becoming increasingly less feasible options for plastic waste disposal, and alternative incineration methods are highly inefficient and unsafe replacements.¹³ This necessitates the development of alternative sustainable waste management systems to replace current end-of-pipe solutions. One promising downstream recycling approach is waste plastic-modified roads. The concept of plastic roads involves repurposing waste plastics by integrating them into road infrastructure using similar methods to traditional paving, but with the key addition of waste plastic. By repurposing plastic waste into pavement, plastic roads introduce a circular plastic economy.

There are two commonly utilized mixing methods in plastic road engineering: the wet process and the dry process. These approaches resemble those used to create widely commercial polymer-modified roads.⁴⁹ The wet process for plastic roads involves mixing waste plastics into liquid asphalt binder, resulting in a waste-plastic modified asphalt binder, which is then laid over quarry aggregates. The dry process entails the incorpora-

tion of waste plastics as plastic-modified dry aggregates. Waste plastic is melted and coated over molten aggregates, and when solidified, forms a dry aggregate road modifier.^{14,15} It is also worth noting the mixed process of plastic road paving, which, although it is not a standardized construction method like the dry and wet process, will be examined in this review. The mixed process combines the mixing methods of the dry process and the lower melting temperature used in the wet process. The plastic partially melts to form a coating around the aggregate, similar to the results of the dry process, except with the use of wet process-affiliated plastics. This process tends to experience less severe storage stability issues than the wet process. See Figure 1.

Currently, multiple states in the United States have implemented plastic road pilot programs, including Missouri, Michigan, California, Hawaii, Pennsylvania, Virginia, and, recently, New York.¹⁶⁻¹⁹ However, these pilot programs are relatively new, and comprehensive life cycle analyses to evaluate their performance and environmental impacts have not been conducted yet. Though some countries in Asia have officially adopted plastic roads as the new pavement standard.²⁰

Plastic roads intend to alleviate the burden on current municipalities and enable safer and more sustainable plastic waste management. This review examines the potential of waste plastic-modified roads as an alternative downstream solution in transitioning to a circular plastic economy in the United States. Key areas of focus include pavement performance and materials properties, economic feasibility, potential environmental impacts, ethical implications, and environmental policy. By synthesizing country-specific and multisectoral factors, this review aims to clarify the developmental scope and long-term viability of plastic road implementation in the United States and evaluate whether it constitutes a viable downstream solution. Additionally, it outlines future areas for research and the development of environmental strategies to advance sustainable plastic waste management strategies.

Figure 1. Diagram of pavement course as well as wet, dry, and mixed process asphalt, where red shaded regions represent plastic-modified content. The dry process contains plastic in the form of dry aggregates, while the wet process contains plastic in its binder, and the mixed method comprises partially coated plastic aggregates and plastic-modified binder.

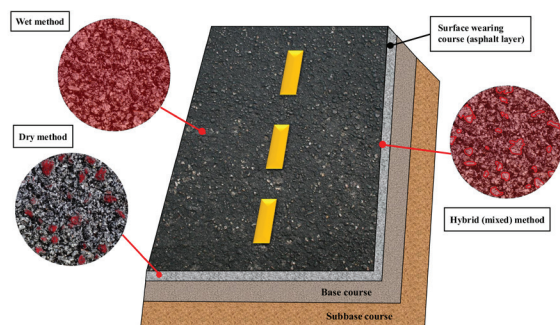


Figure 1: Diagram of pavement course as well as wet, dry, and mixed process asphalt, where red shaded regions represent plastic-modified content. The dry process contains plastic in the form of dry aggregates, while the wet process contains plastic in its binder, and the mixed method comprises partially coated plastic aggregates and plastic-modified binder.

■ Discussion

Environmental Impact:

A potential risk associated with incorporating a persistent environmental pollutant into public infrastructure is the abrasion of microplastics (MP) off the road over time. Although pilot programs are currently observing for signs of this behavior, field studies have not been implemented long enough to conduct a complete life cycle analysis.²⁰ The oldest pilot program in the world in Chennai, India, has only been operational for 21 years; by comparison, asphalt has a typical lifespan of 15-25 years, and modified pavements (e.g., polymer-modified bitumen) extend service life for about 6 additional years.²¹⁻²³ In fact, life expectancy projections for plastic-modified roads are three times longer than those for unmodified asphalt roads.²⁰ Additionally, quantifying MP release from roads in real-world environments remains challenging primarily due to the difficulty in distinguishing road-derived MP from those originating from tire wear, as well as interference from external environmental factors that may reduce the accuracy of MP collection and measurement.²⁴⁻²⁷ Given the limitations in field study data, this paper alternatively reviews two laboratory methodologies that model potential MP release from plastic road samples. Specific studies on potential MP release of plastic roads remain limited, and only in recent years have methodologies to model this phenomenon started to emerge.

Enfrin *et al.* conducted a comparative assessment of the wet and dry processes of plastic road construction, where various plastic types and dosages were observed under environmental stress factors such as duration, water pH, and temperature.²⁸ Post-consumer low-density polyethylene and linear low-density polyethylene (LDPE/LLDPE) and commingled post-consumer polyethylene and polypropylene (PE/PP) were each incorporated into a bitumen matrix via wet process mixing (w) and evaluated at plastic dosages of 1, 2, 4, and 6 weight percent (wt%) of bitumen, respectively. Post-consumer polyethylene terephthalate (PET) and acrylonitrile butadiene styrene (ABS) were combined into the asphalt matrix as plastic-coated aggregates via dry process mixing (d). They were evaluated at concentrations of 0.5, 1, 2, and 4 wt% of asphalt. Methodology was based on wet track concrete abrasion tests, utilizing a wet track abrasion machine to aid in the quantification of MP.

Enfrin *et al.* observed no correlation between MP loss and increasing temperature in ABS (w), where a consistent release rate of $0.43 \pm 0.07 \text{ g/m}^2$ was maintained. On the other hand, lower temperatures facilitated notably high MP release rates in the LDPE/LLDPE (w) sample. At 5°C , the release rate was nearly triple that observed at 25°C . This trend was attributed to potential brittle tendencies of LDPE/LLDPE (w) at low temperatures, which increase susceptibility to fatigue and abrasive mechanical stress.²⁹ This could increase the likelihood of MP shedding, hence the observed result.

MP release from the LDPE/LLDPE (w) sample exhibited no correlation with water pH values, maintaining an average release of $0.53 \pm 0.15 \text{ g/m}^2$ across a pH range of 4 to 10. The absence of a clear release trend under both acidic and alkaline conditions suggests a high resistance to adhesion weaken-

ing. This presents a significant improvement over traditional binders, which are sensitive to moisture susceptibility damage under pH extremes, particularly those caused by acid rain.^{30,31}

Over a duration of 60 minutes, a logarithmic and linear MP release trend was observed for LDPE/LLDPE (w) and ABS (d), respectively. Enfrin *et al.* concluded that this distinction was likely due to the different dispersions of waste plastic in the two methods. In the wet process, the polymer-modified binder is poured over the aggregate layer and creates a thin bitumen film layer, which tends to experience frequent direct contact with the abrasion head at the beginning of the simulation. This physical architecture explains the rapid initial release rate of 0.048 g/m²/min at the 5-minute mark, which then slows down over time, and the wheel abrades through the top bitumen layer and into the asphalt core, where LDPE/LLDPE (w) concentration is lower and more evenly dispersed. In comparison, the plastic content in the dry process is more homogeneously distributed throughout the entire asphalt matrix, resulting in a linear MP release trend.

LDPE/LLDPE (w) initially exceeded ABS (d) by three times more MP released, but was then surpassed by ABS (d) at the 30-minute mark, suggesting differing cumulative release trajectories between the two mix types over the duration of the simulation. It is important to note that this abrasion simulation represents an accelerated, idealized model of surface wear that does not account for the gradual, variable nature of real-world mechanical stress, weathering, or environmental exposure over a road's service life. As such, the crossover trends observed within the 60-minute window should be interpreted as indicative patterns rather than direct predictions of field behavior. Nevertheless, these simulated release trajectories carry implications when considered alongside real-world road lifespans. Road resurfacing typically occurs every 11–26 years, depending on traffic volume and conditions.^{21–23,32} Based on the logarithmic release behavior observed for wet-process roads, MP release would be expected to spike initially following construction or resurfacing when the plastic-concentrated surface layer is most exposed and then gradually decline as the surface weathers. This cycle would resurge with each subsequent resurfacing event. Although dry-process roads introduce approximately 1.5 times more plastic content overall, their uniform dispersion throughout the full matrix depth may moderate MP release over the complete life cycle compared to the wet process.²⁸

Boom *et al.* extended the scope established by Enfrin *et al.* by using a hybrid mixing process of plastic roads.³³ Two asphalt samples were tested, one of which was a combination of LDPE/LLDPE (68/32 wt%, respectively), and the latter consisted of PE/PP. Both samples were prepared using the hybrid mixing process (m). Given the absence of standardized MP release environmental thresholds, the plastic-modified samples were benchmarked against commercially produced polymer-modified asphalt consisting of ethylene-vinyl acetate (EVA) and styrene-butadiene-styrene (SBS).

Since Boom *et al.* employed the MP extraction procedure specified by Enfrin *et al.* (e.g., testing conditions, variables, and materials), results from both studies could be reliably

compared under similar conditions. For comparison and visualization purposes, plastic content data that involved wet process mixing were converted to the same units used for the dry process (wt% of total asphalt mix) and plotted in Figure 2. In wet-mixed asphalt, Enfrin *et al.* reported that bitumen constituted about 5.1 % of the asphalt mix, so this proportion (0.051) was used as the conversion factor.

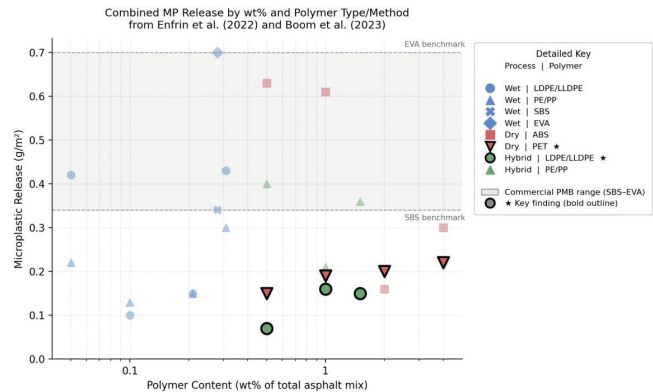


Figure 2: Combined MP release data by plastic content from Enfrin *et al.* and Boom *et al.* (excluding error bars). It reveals a logarithmic trend with increasing plastic content in mixed method LDPE/LLDPE and dry process PET, while other plastic-modified samples exhibit no correlation.^{28,33}

Over a plastic content range from 0.05–0.31 wt% of asphalt, MP release of LDPE/LLDPE (w) and PE/PP (w) exhibited no clear relationship with plastic content, and followed similar high fluctuation patterns, with LLDPE/LDPE (w) having the largest fluctuation range and least consistent results. Enfrin *et al.* and Boom *et al.* observed that at dosage ranges from 0.5%–1.5 wt% of asphalt, mix-processed LDPE/LLDPE (m) released the lowest quantities of MP in the dataset, following a slightly irregular logarithmic trend with MP release rates ranging from 0.07 to 0.16 g/m². A similar trend was observed for PET (d) by Enfrin *et al.*, but with slightly elevated MP release in PET (d), ranging from 0.15 to 0.22 g/m². This logarithmic trend suggests minimal influence on MP release as the plastic content approaches its respective upper dosage thresholds of PET (d) and LDPE/LLDPE (m). Across the overall graphical relationship between plastic content, all recycled plastic modified samples demonstrated lower MP release than the EVA benchmark; however, the samples LDPE/LLDPE (w), ABS (d), and PE/PP (m) exceeded the SBS benchmark at some period. Overall, the two simulations of MP erosion demonstrated that LDPE/LLDPE (m) and PET (d) released the least amounts of MP and exhibited the most evident correlations in MP release trends with plastic content, both following logarithmic curves. Boom *et al.* suggested that this particular trend occurred due to the larger surface area of PET (d) and ABS (d) in contact with bitumen in their respective asphalt when reducing particle size. This may lead to increased adhesive interactions between the binder and polymer-modified aggregate interface compared to other mix configurations. This is beneficial, given that the interfacial adhesion influences the formation of air voids, which in turn affects properties such as water permeability that govern key mechanical characteristics such as abrasion resistance.^{34–37} This phenomenon

was also revealed in the analysis, where MP release of PET (d) and ABS (d) was mainly driven by the aggregate surface area in contact with bitumen, suggesting a positive correlation with the interfacial interactions between plastic additives and bitumen and MP abrasion. Given that PET (d) and LDPE/LLDPE (m) simultaneously released the lowest amounts of MP while also accommodating moderate and upper plastic content thresholds tested in the study, this indicates that more plastic can be recycled as road content per stretch of road while minimizing MP pollution compared to the other mix designs and plastic incorporation methods.

Duan *et al.* investigated LLDPE and recycled polyethylene (PE) pellets modified binder compared to virgin PE powder in a bitumen binder, and quantified MP release in conditions of extreme rutting and moisture.³⁸ 2.5 wt% of bitumen LLDPE pellets, 1 wt% of bitumen PE pellets, and 1 wt% of bitumen virgin PE powder of the asphalt mix's mass were incorporated via the dry process. To simulate MP abrasion, the Hamburg Wheel Tracking Test (HWTT) was employed for the modified asphalt samples in 40,000 wheel-passes, two times greater than the amount established by HWTT testing guidelines. This represents an upper limit of MP release by three orders of magnitude. The particles were filtered and collected through a series of gravitational, density, and chemical separation methods. These simulation results suggested that 1 wt% of bitumen PE pellets (d) and 1 wt% of bitumen virgin powder had a notably higher MP release rate than 2.5 wt% of bitumen LLDPE pellet (d) modified asphalt. Their rates were estimated to be 0.17–0.26%, 0.0674–0.102%, and 0.0355–0.053% of total embedded plastic content, respectively. These data were extrapolated to infer a maximum of 2600 g of lifetime generation of MP and a minimum of 355 g of MP for the highest and lowest release rates in this study, corresponding to 1% PE pellet (d) and 2.5% LLDPE pellet (d) asphalt mixtures, respectively.

From the HWTT results of LLDPE content at 2.5 wt% of asphalt, Duan *et al.* extrapolated that a theoretical 1000 m plastic road with a 10-year service life would have an annual release rate of 0.0053% of the total embedded plastic content. This corresponds to an upper limit of 2209 g of MP/km by a 4-lane interstate road annually, and 920.6 g MP/km/year on a 2-lane local road. To contextualize this degree of MP release, they further compared MP emissions from the plastic-modified asphalt samples to those generated by tire wear. Their findings indicated that the upper limits of road-derived MP emissions were approximately half of those produced by tire abrasion. This suggests that the pollution potential of plastic-modified roads remains substantially lower than that of existing transportation-related sources, particularly given that the simulation overestimated emissions by up to three orders of magnitude compared to projected real-world scenarios.

Once MP are abraded from the road surface, released MP are subject to environmental transport via multiple pathways. Surface runoff during precipitation events is the primary vector, carrying road-derived MP into stormwater drainage networks and ultimately into receiving freshwater and coastal marine environments.²⁰⁰ Aeolian transport of fine MP particles

may also contribute to atmospheric dispersal and deposition in nearby soils and water bodies, though the relative magnitude of wind-driven versus runoff-driven transport remains poorly quantified for road-derived MP specifically.²⁰¹

These studies provide baseline insights into the MP release potential of plastic roads for future work. It can be inferred that MP release potential depends on the engineering properties of the pavement, particularly the bitumen-aggregate adhesion, which contributes to erosion behavior under abrasive and environmental stress. These properties vary with plastic content, construction method, and plastic types. The detailed analysis of which will be discussed in the performance section of this article. For MP release, results and extrapolations from previous studies demonstrate reduced wear-induced MP emission potential in comparison to commercially available polymer-modified bitumen (PMB) roads and virgin polymer-modified roads. Not all dynamic environmental scenarios were modeled; hence, further study and experimentation should be conducted to assess the impact of aging and other factors that cannot be rapidly simulated. There was also a notable lack of environmental thresholds for MP release to interpret the results of these risk evaluation studies. Concrete standards should be established to foster standardized environmental evaluations for risk assessment and decision-makers.

Volatile organic compounds (VOCs) are often emitted during the construction and laying process of traditional hot-mix asphalt (HMA) pavement. This raises concerns about exposing workers to carcinogenic fumes, thereby posing health risks.³⁹ Additionally, observations of non-carcinogenic fumes like polycyclic aromatic hydrocarbons (PAH) have also been reported during the production process at high heating temperatures, particularly the presence of benzo[a]pyrene.^{40,41} Other VOC and PAH observed included benzene, trichloroethylene, tetrachloroethylene, styrene, and dibenz[a,h]anthracene, although those typically constitute low carcinogenic risk.⁴² Notably, Boom *et al.* reported that VOC and PAH emissions increased with higher processing temperatures but decreased with greater plastic content in the asphalt mixture. Variation in carcinogenic and lifetime risk profiles was observed between conventional and plastic-modified binders. Specifically, binders extended with waste plastic exhibited benzo[a]pyrene emissions comparable to or lower than those of C170 and C320 binders, but generally higher levels than the A35P binder.

Beyond processing-stage emissions, the plastic constituents embedded within road surfaces may continuously leach chemical additives, including phthalate plasticizers, brominated flame retardants, and heavy metal-based thermal stabilizers, into adjacent soils and water bodies throughout the pavement service life under thermal cycling and wet-dry exposure. Prolonged UV and oxidative weathering of road-surface plastics can further generate transformation products, such as styrene oligomers from ABS or PS fractions, which may exhibit greater environmental mobility or toxicity than their parent compounds.²⁰²

Plastic tends to become more toxic with increasing recycling cycles due to the accumulation of non-intentionally

added substances like external chemical contaminants.⁴³⁻⁴⁵ This particularly applies to non-recyclable waste plastics like copolymers or chemically degraded plastics, which tend to be more chemically complex due to heterogeneous polymer compositions, additives, or contaminants.⁴⁶ Reclaimed Asphalt Pavement (RAP) is often associated with considerable amounts of PAH and toxic metals, which can form due to incomplete combustion during the RAP recycling process.⁴⁷ Since these compounds are primarily physically dispersed within the asphalt matrix via non-covalent interactions, recent studies have raised concerns about their potential mobilization upon the addition of maltene-based rejuvenators during the renewal process, which may result in chemical leachates contaminating soil and groundwater.^{48,49} Studies on chemical leaching of RAP pavement and stockpiles report inconsistent findings. A study on RAP modified with waste plastic and conventional polymers suggests no apparent groundwater leaching or PAH.⁵⁰ Heavy metals were commonly observed as VOC, though some studies indicated that their quantities did not exceed U.S. Environmental Protection Agency environmental thresholds, and one observed no correlation between plastic content and chemical leaching.⁵¹⁻⁵⁶ Others reported leaching of metals and PAH occurring in excess and suggested that RAP constitutes a high environmental toxicity risk. This included Naphthalene and Dibenzo[a,h]anthracene and metals such as Pb, Zn, Cu, As, Mn, Sb found in both RAP stockpiles and pavements.^{47,57-59}

There are limited findings regarding RAP performance over rejuvenation cycles. Di Mino *et al.* found comparable mechanical properties of an RAP asphalt sample modified with recycled hard plastics and graphene nanomaterial across three rejuvenation cycles. Additionally, each aging cycle required similar dosages of rejuvenator, suggesting no significant performance degradation with increasing rejuvenation cycles over the course of three life cycles. However, there remain uncertainties in the mechanical performance of RAP mixtures as studies present controversial findings.^{59,60}

Significant gaps remain regarding the physical and long-term limitations of RAP, particularly in the context of its circularity. While recent studies have performed multiple life cycle analyses with RAP, the long-term feasibility of continuous rejuvenation remains uncertain without prolonged field assessments. If the number of viable rejuvenation cycles constrains RAP implementation, the eventual disposal of pavement materials becomes problematic. Even with assumptions of minimal chemical leaching, these materials may still pose hazardous risks and face the same challenges as current disposal methods, particularly the safe management of large volumes of chemically contaminated waste. The main challenge involves the handling of persistent, toxic compounds, which often requires active physical intervention and deliberate chemical stabilization strategies. Additionally, the environmental toxicity of plastic-modified pavements during their service life remains uncertain and entails prolonged and comprehensive life-cycle assessments of environmental pollutants, including chemical leachates and microplastics.^{61,62}

Performance:

Polymers are widely commercialized modifiers used in pavement engineering. Their use has been noted to improve asphalt durability by enhancing resistance to thermal stress and extending pavement lifetimes, resulting in higher-performing roads.^{63,64} Still, there remains potential for improved mechanical performance, environmental contribution, and greater recycling applications; hence, the exploration of circular plastic modifiers.

Commonly tested post-consumer plastics and polymers for road modification include low-density polyethylene (LDPE), high-density polyethylene (HDPE), polypropylene (PP), polyvinyl chloride (PVC), polystyrene (PS), polyurethane (PU), ABS, and PET. Key performance metrics of asphalt include rheology, moisture susceptibility and aging, viscoelasticity, and morphological and storage stability. There are three different types of polymeric modifiers: plastic, thermoplastic elastomer, and reactive polymer. Inert plastics and thermoplastic elastomers are combined into asphalt via physical mixing.⁶⁵ The inert polymers are physically mixed and dispersed into a bitumen matrix to achieve a homogeneous state, which is maintained via preferential, van der Waals interactions and adsorption between bitumen and aggregate components.^{15,66,67} However, since the polymer-bitumen interface is not chemically bonded, storage stability issues for polymer components dissociating from the mix at elevated temperatures have been frequently reported.^{68,69} Casey *et al.* observed this phenomenon in PP-modified bitumen, where the PP migrated to the surface during high-temperature storage, and formed a film layer of pure PP at the bitumen surface that could not be re-integrated into the asphalt. This behavior compromises the durability and service life of the pavement and may inhibit wear-induced MP erosion as previously discussed.^{70,71} Reactive polymers, on the other hand, undergo chemical modification, in which prepolymers are reacted with the asphalt components to form chemical structures composed of polymer and asphalt constituents. This modification method tends to improve high-temperature storage stability due to the chemical bonding between the polymer and asphalt molecules, making them less vulnerable to direct physical changes.^{63,71,72,73}

Table 1: Summary of key points regarding pavement performance advantages and disadvantages of LDPE-, HDPE-, PP-, PET-, PVC-, ABS-, PU-, PC-based waste plastic modified asphalt.

Plastic Type	Observed Performance Key Points	Processing and Structural Characteristics
LDPE/LDPE	- Improves moisture susceptibility and high-temperature deformation resistance.^{15,74-79} - Often increases viscosity and elastic recovery, but can suffer phase separation at higher dosages.^{26,70,80,89,91,93}	- Low average melting point (105-120 °C)^{15,74,78} - Suitable for wet-mixing methodologies.^{15,74-76} - Forms a 3D plastic-asphalt network via molecular chain entanglement.^{15,82-86} - Requires chemical stabilization (e.g., polyphosphoric acid or sulfur cross-linking) to reduce phase dissociation.^{86,91,93}
HDPE	- Improves the moisture susceptibility and viscosity, but storage stability and elastic recovery are inconsistent and may require additives.^{15,68,76,88,89,94,100}	- Rigid, dense homopolymer with high molecular weight that promotes viscous behavior.^{15,84-86} - Melting point is ~135 °C; recommended for wet process mixing.^{15,84-86} - Stability issues are minimized at a 4 wt% dosage.⁶⁸ - Polyphosphoric acid at 0.8 wt% minimizes component dissociation and increases viscosity.⁸⁸
PP	- Increases asphalt stiffness, rutting resistance, tensile strength, and fatigue life.^{68,106-109} - Micronized PP (optimally 177 µm, 55 min curing) improves hydrophobicity by up to 41.81% increase in contact angle, while storage stability is dosage-sensitive.^{87,112}	- Thermoplastic with wide-ranging molecular weight and a melting point between 160-170 °C.^{15,96,104-106} - Well-suited for wet process.^{15,96,104-106} - Incorporating elastomeric additives (crumb rubber, SBS, or SBR) significantly remedies stability, moisture, and rutting deficiencies.^{113,114}
PS	- Generally increases stiffness and viscosity.^{115,130} - Can worsen low-temperature brittleness, compactibility, adhesion, and abrasion resistance.^{115,130}	- Semicrystalline isotactic polystyrene possesses a melting range between 264-272 °C.¹³⁹ - High melting points (e.g., in high-melting PS) hinder the formation of a stable, covalently bonded polymer-asphalt network during wet process mixing.¹²¹

PET	- Improves stiffness, viscosity, elastic recovery (e.g., 29% rut depth reduction and 16% increase in elastic recovery at 4 wt%) and high-temperature rutting resistance.^{84,131,132,145} - Conflicting findings on adhesion and moisture resistance, potentially increasing air voids and raising structural vulnerability.^{146,128,132,137,141}	- Inert thermoplastic with very high molecular weight and a melting point around 250–265 °C.^{131–133} - Processing temperature significantly exceeds average bitumen mixing limits, rendering it unsuitable for conventional wet process mixing according to some studies.^{13,68,135}
PVC	- Increases stiffness, viscosity, rutting resistance, and elastic recovery, while also improving adhesion and moisture resistance.^{146–157} - High melting points can disrupt homogeneous blending during wet process mixing.⁶⁸ - Studies observed threshold effect in air void content: it initially decreases with dosage but then increases significantly at elevated concentrations.^{158,157} - Generates highly toxic, corrosive HCl gas, benzene, and halogenated hydrocarbons at elevated processing temperatures.^{159–161}	- Highly crystalline polymer with a melting point of 100–260 °C aligning with or exceeding crumb rubber-modified Hot Mix Asphalt temperatures (166–204 °C).^{154,144,146} - Applicable for both wet and dry processes.^{144,146} - Chlorine gas mobilization requires mitigation techniques. Ragab <i>et al.</i> suggests absorption via di-2-ethylhexylphthalate.¹⁶²
ABS	- Usually improves stiffness, viscosity, high-temperature deformation resistance, and sometimes elastic properties.^{164,122,124,166,167} - High melting point variants of ABS introduce severe wet-mixing incompatibility.⁶⁸ - Highly variable composition and chemical contamination within consumer waste streams result in unpredictable mechanical properties.^{160–165} - High dosages (e.g., 9 wt%) degrade rheological and fatigue performance due to excessive stiffness and low-temperature brittleness.^{124,167,168} - Releases hazardous VOC and toxic benzene byproducts via styrene volatilization at processing temperatures (140–180 °C).^{164,170–172}	- Heterogeneous, multiphase amorphous terpolymer with a glass transition temperature around 100 °C but no definitive melting point.^{42,163} - Chemical degradation initiates at 120 °C; major benzene emissions occur above 140 °C.¹⁶⁴ - Recommended maximum dosage threshold is restricted to 4–5 wt% of bitumen.^{124,167,168}
PU	- Recycled PU particles enhance low-temperature cracking resistance and aging resistance.¹⁷⁸ - Recycled PU improved rutting resistance relative to SBS-modified PMB and enhanced low-temperature cracking and aging resistance; thermal decomposition products were additionally found to improve rutting resistance.¹⁷⁸ - Reactive PU yields better homogeneity and storage stability than recycled PU alone due to the formation of new chemical structures rather than physical dispersion.^{67,69,72}	- Storage stability deficiencies in recycled PU can be mitigated and improved by 42.3% through the inclusion of SBS elastomeric additives.¹⁷⁸ - For reactive prepolymer PU, achieves stability through active chemical cross-linking rather than weak, physical van der Waals physical dispersion.^{67,69}

LDPE:

LDPE is a frequently assessed waste thermoplastic pavement modifier. It is a high-molecular-weight homopolymer, characterized by its long, linear, and flexible polyethylene chain. Compared to denser forms of polyethylene, it has a lower strength and temperature resistance, which makes it suitable for consumer-based applications in single-use packaging and containers. Commercial LDPE is most often incorporated into pavement via wet-mixing due to its compatible low average melting point of 105–120 °C.^{15,74–76}

The most notable and consistently observed improvements associated with LDPE-modified asphalt are moisture susceptibility and high-temperature deformation resistance.^{77–79} On the other hand, low temperature performance and stiffness remain variable. Studies show increased stiffness at lower temperatures; however, some have revealed those effects to be excessive, noting retained stiffness and flow resistance even at warmer temperatures. This concerns pavement flexibility and susceptibility to brittle fracturing.^{26,78,80,81}

Several studies attributed an increased viscosity compared to unmodified bitumen, and in certain cases, gel-like behavior with the addition of LDPE.^{82–84} This phenomenon is likely due to increased polymer chain entanglement, resulting from swelling of the plastic components and subsequent formation of a 3D plastic-asphalt network.^{15,85,86} Elasticity is similarly affected by this molecular-level physical crosslinking, though research results are not entirely aligned. Punith *et al.* reported a poor elasticity with the sole use of LDPE; several other studies noted an opposite, beneficial effect.⁸⁷ Notably, Zoorob and Suparma and Dalhat *et al.* consistently reported a 14% increase in elastic recovery, while the latter study also showed a positive correlation between LDPE content (2–8 wt% of bitumen) and elastic performance.^{78,88} Ho *et al.* suggested that the variations among different studies can be primarily caused by varied molecular weight and molecular weight distributions, which influence the degree of entanglement.⁸⁹ Other factors related to chemical structure, such as polarity and crystallinity, also have pronounced effects on the compatibility between

polymer and asphalt components, which governs related viscoelastic and rheological properties.^{15,85,90} LDPE that has a broad molecular weight distribution exhibited higher viscosity with increasing plastic content, while an inverse relationship was observed in PE wax that had a much narrower weight distribution. Compared to base asphalt, LDPE-modified binder exhibited higher viscosity overall with high-molecular-weight LDPE, in particular, harboring the highest viscosity to its low-weight counterpart.

Due to the nature of the physical modification process, phase separation issues were observed as plastic content dosage increased to 4 wt% asphalt.⁹¹ Similarly, Naskar *et al.* noted poor stability and phase separation issues at 7 wt% of bitumen.⁹²

Therefore, additive methods have been explored to enhance the performance of LDPE-modified asphalt further. Polyphosphoric acid (PPA) (0.8 wt% bitumen) with a dosage of 4 wt% bitumen LDPE was found to minimize dissociation within the mix.⁶⁸ Other stability-enhancing methods include chemical cross-linking with sulfur agents, which have been shown to improve the dispersion of the inert plastic components throughout the asphalt matrix. Despite improvement in the relative stability of waste-modified bitumen, such methods currently do not outperform commercial PMB.⁹³

HDPE:

HDPE is a rigid and denser form of polyethylene with a typically high molecular weight. HDPE has a higher melting point than LDPE (about 135 °C), which is just below the average mixing temperature of bitumen.^{94,95} Thus, it is generally recommended for wet process mixing.^{15,96}

Several studies have demonstrated that incorporating post-consumer HDPE into asphalt binder can improve the moisture susceptibility of pavement compared to unmodified specimens. Köfteci (2016) reported that an HDPE additive at concentrations of 1, 2, 3, and 4 wt% of bitumen enhanced resistance to water damage, as evaluated with AASHTO T283 testing specifications, compared to standard pavements.⁹⁷ This improvement was attributed to increased storage stability, which is positively correlated with reduced moisture susceptibility. The inclusion of HDPE has also been associated with increased binder viscosity compared to unmodified bitumen, as indicated by elevated softening points and bitumen stiffness.^{76,98} Additionally, molecular weight plays a significant role in a fluid's viscosity, potentially explaining the increased viscous behavior observed with the use of HDPE due to its high molecular weight.⁹⁹

Inconsistency remains for the stability behavior of HDPE-modified bitumen across studies. While some studies have reported enhanced stability in correlation with increased moisture resistance relative to unmodified bitumen, others have observed adverse effects.^{89,97,100,101} Poor storage stability was noted by Casey *et al.* and Appiah *et al.*, where the HDPE modifiers propagated inconsistencies in mechanical properties and durability under thermal stresses in comparison to proprietary PMB and unmodified bitumen, respectively.^{68,100} Furthermore, Appiah *et al.* compared HDPE with PP and associated a disproportionate and fluctuating effect on rheological and vis-

coelastic properties with the incomplete blending of HDPE, thus concluding that HDPE is unsuitable for pavement applications. Similarly, Casey *et al.* found HDPE to be feasible only at a dosage of 4 wt% of bitumen, where such stability issues were minimized. As a complementary approach, Casey *et al.* observed that the use of 0.8 wt% of bitumen PPA as a chemical additive minimized dissociation between HDPE and binder components, as well as increased overall binder viscosity.

Deformation resistance at high temperatures also varies across studies. While several reports note enhanced rutting resistance, others have found it to be comparatively lower than that of commercial PMB, such as those incorporating EVA and SBS; nevertheless, it still adhered to standard deformation thresholds.^{68,88,101-103}

There are limited reports regarding elasticity. Similar to findings by Costa *et al.* and Dalhat *et al.*, who found that HDPE modifiers alone could not meet the AASHTO TP 70 elastic recovery standard for performance grade asphalt without the addition of an elastomeric polymer additive.^{88,103}

PP:

PP is a thermoplastic with a wide-ranging molecular weight and a melting point between 160–170 °C, making it commonly suited for the wet process mixing.^{15,96,104-106}

Regarding pavement rheological performance, studies frequently note increased asphalt stiffness and improvements in fatigue life compared to unmodified bitumen, implied by softening points and decreased penetration value, and enhanced tensile strength with various post-consumer PP modifiers.^{68,106-108} Similarly, high temperature deformation resistance is also reported to improve. The same studies indicated an improved resistance to permanent deformation (e.g., rutting) and greater stiffness. Additionally, Tapkin noted an extended fatigue life at 1% PP by 27% compared to the unmodified bitumen.¹⁰⁹ However, these findings aren't entirely conclusive regarding overall durability. Some studies report an increase in abrasion loss and variable stiffness characteristics, where stiffness rapidly increases with plastic dosage yet exhibits poorer damping capacity compared to unmodified bitumen.¹¹⁰ Likewise, studies have also attributed a reduced resilient modulus and ductility, which could result in poorer elastic recovery and low temperature cracking, respectively.^{106,111} A possible explanation for the observed reductions in fatigue and deformation resistance is reduced impact absorption and excessive brittleness, compounded by poor moisture resistance. This was evident in the experimental findings of Ahmedzade *et al.*, where increased viscosity and stiffness were accompanied by a decrease in ductility, indicating an excessive stiffening effect that may contribute to brittleness and heightened fatigue susceptibility.^{106,110} However, other studies report contrasting trends. Dalhat found that micronized PP could be used to achieve super-hydrophobicity of the pavement.¹¹² Moisture resistance was primarily influenced by particle size and curing time, with an optimal condition achieved at a particle size of 177 µm and a curing time of 55 minutes. Moisture resistance was significantly enhanced under these conditions, noted by a

41.81% increase in contact angle compared to an unmodified bitumen binder.

The viscoelastic behavior of PP-modified asphalt mixtures remains insufficiently characterized, with limited and variable findings reported in the literature. Otuoze *et al.* consistently observed a deterioration in viscoelastic and rheological properties when PP was incorporated into asphalt binders, while Dalhat *et al.* reported no clear trend or adherence with PP in relation to elastic recovery adhering to AASHTO TP 70 elastic recovery standards.^{88,111} Without elastomeric polymer additives, PP-modified binders consistently exhibited deteriorated elastic recovery values below 20. The initially poor elasticity associated with solely PP modifiers was attributed to the incompatibility of PP with the asphalt components and the resulting poor storage stability.⁸⁷ The same study found that stability could only be maintained at dosages up to 2 wt% of the bitumen binder; higher concentrations required the inclusion of an elastomeric polymer. Separately, Chiang *et al.* and Vamegh *et al.* supported this conclusion, noting significant improvements in thermal stability, as well as moisture and rutting resistance when using elastomeric polymer additives (crumb rubber and SBS in the former, and synthetic styrene-butadiene rubber (SBR) in the latter), opposed to solely PP-modified bitumen.^{113,114}

In contrast, other studies have reported that the sole use of PP without additives had a profound effect on the bitumen's homogeneity and storage stability at an optimum dosage of 3 wt% of bitumen.⁹⁶ Appiah *et al.* attributed this to the lower molecular weight and polarity of PP, which supports earlier observations by Ho *et al.* linking lower molecular weight polymer modifiers to higher storage stability.⁸⁹ Furthermore, Appiah *et al.* found PP more favorable compared to HDPE due to its slight, linear improvements to the viscosity, deformation properties, and storage stability compared to the inconsistent mechanistic qualities of HDPE.¹⁰⁰

PS:

Expanded polystyrene (EPS) foam is banned in at least twelve states and numerous local jurisdictions due to concerns over its environmental persistence, limited recyclability, and potential health risks.¹¹⁵⁻¹¹⁹ Post-consumer PS has a wide melting range depending on its state, though the melting point of semicrystalline isotactic polystyrene is between 264 and 272 °C.¹²⁰

Generally, PS has been reported to increase the stiffness and viscosity of asphalt samples compared to unmodified specimens, given its increased softening points and enhanced rotational viscosity.¹²⁰⁻¹²⁴ Though thermal deterioration has been documented at both high and low temperature extremes. Hasan *et al.* noted PS-modified bitumen to have a faster stiffening rate at low temperatures in comparison to unmodified bitumen, which may lead to excessive rigidity and increase susceptibility to brittle fractures, and such susceptibility in low temperatures has been noted in other studies.^{122,123}

PS has also been associated with increased air void content in asphalt mixtures compared to traditional stone aggregate asphalt. Using the dry process, Vasudevan *et al.* compared re-

cycled polystyrene, polyethylene, polypropylene, laminated plastics, biaxially oriented polypropylene, and unmodified stone aggregates.¹²⁶ They found that the recycled polystyrene sample showed the poorest compactibility, which was attributed to the relatively high air void content due to the weak adhesion between the polymer components and binder. Such characteristics may induce structural issues such as asphalt durability, moisture susceptibility, fatigue, and deformation. Additionally, wet-mixed applications face similar compatibility issues due to the high melting points of certain polystyrene types, particularly high-impact polystyrene. Hasan *et al.* reported that the high melting point hindered the formation of a covalently bonded polymer-asphalt network, which may prevent a stable and homogeneous PS-modified asphalt formulation.¹²² On the contrary, studies testing lower melting point PS, namely EPS with acetone and ethyl acetate solvents, found adhesion to be enhanced compared to unmodified bitumen.¹²⁷

Vila-Cortavitate *et al.* reported increased air voids and higher Cantabro loss at 1 wt% of asphalt PS content, indicating reduced abrasion resistance compared to the unmodified bitumen.¹²³ Overall, these effects may stem from poor moisture resistance, elevated air void content, or weak aggregate-binder interactions.¹²⁸ Despite these limitations, the same study reported unusually improved high-temperature deformation resistance. These discrepancies introduce uncertainties regarding the thermal susceptibility and overall consistency of PS in pavement applications. Moreover, it is a problematic and ecotoxic material that is banned in several states due to its persistence in the environment and overall health hazards.^{129,130} Integrating PS into downstream recycling efforts like plastic roads could indirectly enable its continued production and use, which would be counterproductive given the non-biodegradable nature of polystyrene, and could potentially contribute to further environmental damage.

PET:

PET is an inert thermoplastic polymer with a very high molecular weight and melting point at around 250–265 °C.^{131–134} This high melting point significantly exceeds the average bitumen mixing temperature and is deemed by some studies as unsuitable for the wet process mixing.^{15,68,135}

Guru *et al.* and Ghabchi *et al.* suggested a strengthening effect on adhesion at the polymer-bitumen interface, which is an overall positive indicator towards fatigue, moisture susceptibility, and deformation resistance.^{131,136} This effect was confirmed in dry process implementation with Quartzite and Granite aggregates by Ghabchi *et al.*, who found improved stripping resistance to unmodified bitumen, which is primarily affected by the water resistance abilities of the asphalt. On the other hand, Ahmadiania *et al.* noted that the crystallization of the PET, which trapped the sticky binder around the aggregate surface and reduced the surface area in contact with the asphalt film, thereby weakening the aggregate-bitumen adhesion.¹³² This raises concerns about air void potential and moisture susceptibility, primarily due to the direct influence of aggregate-binder adhesion.^{36,128} Such vulnerability has been acknowledged by several studies in findings of increased air void content com-

pared to the unmodified control asphalt, though a study in India found that the air voids still adhered to country specifications.^{137–141} Overall, the possibility of weakened adhesion and air voids necessitates further research on the crystallization behavior of PET in the binder mix, as it can catalyze adverse effects on performance under certain conditions.

Rheology and fatigue remain mostly unaffected by the contradictory moisture susceptibility behavior derived from aggregate-asphalt adhesive behaviors, and in fact, are noted to improve overall compared to unmodified bitumen.^{132,138,142,143} As such, studies have documented improvements in high temperature deformation resistance, stiffness, and viscosity compared to unmodified asphalt mixtures.^{86,131,142} In particular, Ahmadiania *et al.* noted a 29% reduction in rut depth and 16% increase in elastic recovery at 4 wt% of bitumen PET content in comparison to a conventional bitumen mix.¹³² These effects are likely attributed to the enhanced stiffness and elasticity of the asphalt, leading to enhanced recovery capacity from deformation than unmodified bitumen binders.^{85,139,140} Mashaan *et al.* suggested that the greater elastic recovery rate was possibly due to the compatible swelling of the PET within the asphalt.⁸⁶

PVC:

PVC is a thermoplastic polymer produced from its vinyl chloride monomer, with a high molecular weight and crystallinity.¹⁴⁴ It is the world's third most produced synthetic polymer and versatile in creating both rigid and flexible materials, often as single-use packaging or piping and construction appliances.¹⁴⁵ PVC has a crystalline melting point of 100 – 260 °C, which falls level with or slightly above average crumb rubber-modified Hot Mix Asphalt mixing temperature (166 – 204 °C), thus making it suitable for both the wet and dry processes.^{134,144,146} Though PVC is often incorporated via the wet process, some studies found the high melting point to hinder mixing and perform unsuitably when utilizing the wet process.⁶⁸

Post-consumer PVC-modified asphalt is commonly associated with greater stiffness and viscosity, observed with increased softening points and lower penetration values compared to neat specimens.^{146,147} In addition, this modifier enhances resistance to rheological and fatigue stress, such as resistance to permanent rutting deformation.^{148,149–152} Further evidence of an extended low-temperature fatigue life is suggested by the higher asphalt stiffness relative to unmodified bitumen, which is a typical correlation.^{147,153,154} Since asphalt is expected to perform under a wide range of temperatures, an optimal formulation requires a balance between stiffness and flexibility in order to resist both high temperature deformation and low temperature cracking. In the case of a stiffer asphalt mix, excessive rigidity may increase brittleness and compromise performance at low temperatures.¹⁵⁵ Fakhri *et al.* observed both resistance to brittle fractures and enhanced stiffness compared to conventional bitumen.¹⁴⁷ This study suggests that it is possible to achieve enhanced asphalt stiffness without risking brittle susceptibility, suggesting that PVC-modified asphalts can potentially achieve balanced deformation characteristics.

PVC is also found to improve elastic recovery relative to unmodified asphalt.¹⁴⁷ This is consistent with reports of enhanced rutting resistance, which correlates to the elastic recovery ability of pavement to return to its original form after deformation. Additionally, PVC has been reported to enhance adhesion between the bitumen-aggregate interface, which minimizes air voids and susceptibility to moisture damage.^{147,156} Rasel *et al.* and Rahman *et al.* affirmed that the air void content of the PVC-modified asphalt sample remained within limits specified by The Asphalt Institute.^{156,157} In addition, both studies revealed similar patterns in air void content levels, in which it initially decreased with increasing PVC dosage but later increased as the dosage reached higher levels. Overall, there appears to be an effective PVC dosage threshold for optimizing adhesive strength and related properties.

PVC has a high hazardous emission risk when heated to elevated temperatures, specifically, the release of large quantities of HCl gas and potentially other hazardous gases such as benzene and halogenated hydrocarbons.^{158,159,160} HCl gas is highly toxic to the environment due to its corrosive acidic nature, which is a major contributor to ocean and soil acidification.¹⁶¹ In the context of pavement applications, which are routinely subjected to variable thermal and environmental stresses, the emission potential of PVC presents a significant challenge to mitigate and control over the pavement's service life.^{152,158} To mitigate the mobilization of such toxic compounds, studies have proposed methods to absorb chlorine content during asphalt preparation using di-2-ethylhexyl phthalate.¹⁶²

ABS:

ABS is a heterogeneous, multiphase thermoplastic terpolymer consisting of acrylonitrile, butadiene, and styrene.¹⁶³ ABS can be incorporated into asphalt mixtures via both dry and wet processes, owing to its amorphous structure. Although it lacks a definitive melting point, ABS exhibits a glass transition temperature around 100 °C.⁴² Notably, ABS is documented to experience chemical degradation at temperatures starting from 120 °C and significant benzene emissions at temperatures exceeding 140 °C. This is attributed to the volatilization of the styrene phase. The heterogeneous composition of ABS introduces variation in the proportions of its constituent polymers, which may confer a broad range of thermal and mechanical properties.¹⁶⁴ Indeed, this variance was reflected in the melting points of ABS, where Casey *et al.* noticed mixing incompatibility issues between the high melting point of the ABS that hindered conventional wet process mixing.⁶⁸ In contrast, other studies expressed no issues. ABS and other waste-derived heterogeneous copolymers have significantly variable material composition, given that consumer waste streams encompass different variants and compositions of ABS plastic and potential chemical contamination. This can lead to different mechanical properties in the resulting pavements, given that mechanical behavior is highly dependent on the chemical compositions and microstructure of the polymer additive.¹⁶⁵

Rheological and fatigue studies demonstrate a modest degree of variability, with most reporting an increase in stiffness compared to unmodified binders, suggesting higher viscosity

and deformation resistance in high temperatures, though not superior to certain commercial PMB.^{104,122,124,166,167} Aligned with these findings, Singh *et al.* and Fonseca *et al.* also noted an increase in elastic properties. In contrast, Morales *et al.* observed degraded rheology and fatigue performance compared to unmodified bitumen at an ABS concentration of 9 wt% of bitumen. They attributed this to excessive stiffness that may have led to low-temperature brittleness. Additionally, ABS was evidenced to have little to no effect on viscosity. Slight upward increments were attributed to the oxidation of bitumen during processing. It is worth noting that other studies recommended against exceeding ABS dosages of 4–5 wt% of bitumen. The higher dosage of 9 wt% of bitumen employed by Morales *et al.* may have contributed to the negative performance effects observed.^{124,167,168} Slight decreases have been observed in the air voids of ABS-modified asphalt compared to unmodified bitumen.¹⁶⁶ Audy *et al.* reported similar findings in comparison to PET via dry process modification, as ABS aggregates were observed to have better adhesion with the bitumen compared to PET, which would lead to decreased air void content.¹⁶⁹

Finally, during its additive manufacturing applications, ABS is known to emit VOCs harmful to humans when heated to high temperatures ranging from 170 to 240 °C.^{170–172} Tiganis *et al.* documented the mobilization of benzene byproducts via the volatilization of styrene when ABS was heated to temperatures higher than 140 °C.¹⁶⁴ The asphalt processing temperatures utilized in the study (140–180 °C) approximated industry standard wet process mixing thresholds, which indicates credible potential for harmful emissions during the asphalt mixing process.^{42,134}

PU:

PU differs from other aforementioned polymers, like polyethylenes or PS, in that it is a chain of alternating copolymers. There are various forms of PU due to its versatile chemical composition, resulting in several different applications such as rigid forms, foams, adhesives, coatings, fibers, etc.^{73,173,174}

Unlike the inert thermoplastic polymers previously discussed, most PU-modified asphalts utilize active PU, the most common being prepolymers. Prepolymers are reactive polymers incorporated into pavement via chemical modification.⁷² Unlike physical modification, in which the polymer is physically dispersed throughout the asphalt mix and maintains stability primarily through van der Waals interactions, the prepolymer PU is an active component that reacts with asphalt constituents to form new chemical structures and functional groups, achieving a more stable homogeneous state. This results in a higher performing and uniform mixture at high temperatures, since the polymer–asphalt interface is reinforced by chemical bonding.^{67,69}

Overall, PU is generally found to enhance the viscoelastic and rheological properties of asphalt mixtures compared to conventional types.¹⁷⁵ Wei *et al.* incorporated PU prepolymer at dosages of 3, 5, and 7 wt% of the total asphalt mix. They observed increased complex shear modulus values compared to conventional asphalt mix, indicating greater stiffness and deformation resistance.¹⁷⁶ Although liquid PU at 4.1 wt% of asphalt

yielded rutting resistance comparable to SBS-modified mixtures, it still outperformed base asphalt.¹⁷⁷ Similarly, Hao *et al.* found that recycled PU improved rutting resistance compared to an SBS-modified PMB sample and also enhanced low-temperature crack resistance and aging resistance.¹⁷⁸ Chemical analysis revealed that the aging process induced a decomposition reaction of the recycled PU, resulting in broken-down molecular products that improved the rutting resistance. Overall, the mechanical properties of recycled PU similarly reflect those enhancements produced by virgin or prepolymer PU in comparison to unmodified bitumen. Thermal susceptibility is improved, and mechanical properties are well-balanced so that the pavement is neither overly rigid nor pliable.

The main difference between recycled PU and reactive PU-modified asphalts is storage stability. Unlike reactive PU prepolymers, adding recycled PU particles typically results in reduced thermal storage stability.^{175,178} This reduction is likely due to differences in material modification methods and the distinct interactions between polymer and asphalt constituents when comparing prepolymers, waste thermoplastics, and thermosetting elastomers. Poor stability may be mitigated through the incorporation of elastomeric additives; for example, Hao *et al.* demonstrated that storage stability was improved by 42.3% when using SBS compared to formulations without elastomer enhancement.¹⁷⁸

Economic Feasibility:

The long-term economic model of plastic-modified pavements is a critical point of evaluation in the context of public infrastructure renewal and municipal budget constraints. Current pavement networks in the United States contain asphalt composed mostly of unmodified bitumen or PMB mixed with quarry aggregates.¹⁷⁹ Partial substitution of these materials with post-consumer plastics introduces elements of a circular economy while reducing raw material costs.¹³⁷ Multiple economic analyses demonstrated that waste plastic modified bitumen can match or even reduce the costs of construction and maintenance compared to conventional bitumen. For example, Yao *et al.* reported a cost reduction of up to 31.8% compared to virgin PMB construction.¹⁸⁰ Singh *et al.* noted a 4.3 and 9.5% reduction in surfacing costs when comparing LDPE to crumb rubber PMB. The study observed that laying costs similarly reflected those typically of Viscosity Grade 40 bitumen.¹⁸¹ The same evaluation by Yao *et al.* demonstrated extended pavement lifetime with waste plastic integrated asphalt matrices in comparison to unmodified or PMB pavement, thus resulting in lower maintenance frequency and cost.¹⁸⁰ This was noted by a total life cycle cost reduction ranging from 14.5 to 26.2%. Aside from economic benefits, the reconstruction and installation of new roads can also facilitate socioeconomic growth. Studies assert that under the right economic circumstances and spatial factors, the construction of transportation infrastructure often promotes increased economic productivity and competitiveness.¹⁸²⁻¹⁸⁶ According to American socioeconomic studies, there is a causal relationship between income inequality and a lack of domestic transportation.¹⁸⁷ Expanding road infrastructure construction initiatives, particularly sustainable road

technologies like plastic roads, could therefore play a critical role in reducing spatial inequality and income disparities in marginalized geographical regions.

Government investment towards road infrastructure restoration has substantially grown in recent years, with recent enactments of road investment policies such as the Infrastructure Investment and Jobs Act passed in 2021.^{188,189} Results of such efforts are illustrated by the 2025 report from the American Society of Civil Engineers, which indicated a record high performance of U.S. roads.¹⁸⁸ Overall, the economic feasibility of plastic roads and current upward road investment trends and attention from stakeholders suggest high potential for the implementation of plastic roads on the national level, that is, until they have undergone comprehensive risk and lifecycle assessments.

Policy:

The shifting U.S. regulatory landscape is predicted to have considerable impacts on the employment of waste plastic recycled pavement technology. Currently, waste-plastic-modified roads are being constructed by private companies, such as Dow Transportation and the UK's MacRebur, which state departments of transportation and smaller administrative bodies hire.^{190,191} Plastic roads have not been implemented on a larger scale primarily because 1) they are a new technology and still undergoing laboratory tests, and 2) past infrastructural regulations required technology to adhere to a comprehensive environmental review.¹⁹² However, shifting infrastructural policies allude to changes that may enable either a potentially more streamlined or complicated process for the expansion of the plastic road industry. In 2025, several environmental and permitting evaluation processes in the context of transportation infrastructure were rescinded, including greenhouse gas regulations and the removal of the authority of the Council on Environmental Quality to oversee the National Environmental Policy Act. The intention was to streamline permitting processes to expedite road infrastructure construction projects, though some experts state that these rescissions would in fact have the opposite effect on energy and transportation sectors, claiming that the absence of environmental standards without the oversight of the Council on Environmental Quality would lead to increased process and production delays.¹⁹³⁻¹⁹⁶ These policies transcend into the plastic road industry as well, which relies on environmental benchmarks to conduct environmental toxicity evaluations, as emphasized by previous environmental assessments.^{28,33}

Shifting environmental and infrastructural policies are key components in the scalability potential of the plastic road industry. The absence and further retraction of environmental standards will create more complicated risk evaluations during the current experimental stages of laboratory and pilot studies of plastic roads.

Ethics:

While plastic roads facilitate a downstream solution to alleviate the burden on current waste management systems, there are several ethical questions about their long-term sustainabil-

ity and impact on producers, localities, and current upcycling streams.

Failing downstream waste management systems already demonstrate the insufficient long-term sustainability of downstream solutions. Sole reliance on end-of-pipe solutions obscures the vast accumulation of plastic waste and foregoes attaching accountability to more inherent issues in materials regulation, pollution, and unethical producers. The concept of waste plastic-modified pavements should adhere to a more circular economic model rather than traditional end-of-pipe principles, which burden consumers and municipalities with issues created by producers.^{197,198} Although plastic roads theoretically offer a more circular model through RAP, the environmental risks in this practice remain uncertain.^{50,52,53}

It should be clarified that waste disposal systems like plastic roads are not completely ineffective; rather, a short-term remedy is critical to resolving current waste municipalities' burdens and preventing further plastic accumulation in ecosystems. However, there needs to be greater emphasis on upstream solutions to promote the development of a circular plastic economy and phase out traditional linear end-of-life plastic life cycles. Ethical motivators for future research should explore whether plastic roads can strengthen existing downstream waste management systems with a less ecotoxic, socially ethical, and sustainable model.

■ Conclusion

Today, plastic roads are underdeveloped in terms of performance, environmental impact, circularity, and regulation. Though plastic roads have proven to be more economical and certain compositions exhibit mechanical advantages, there still remain significant process gaps and disadvantages, such as worsened storage stability due to poor adhesion between asphalt and plastic constituents. The inconsistent material composition and chemical impurities of post-consumer stream plastics also introduce microstructural and chemical inconsistencies that can alter the mechanical properties of the final modified pavement. This applies, in particular, to heterogeneous copolymers, which retain a wider range of polymer structural variations. Thus, future research should consider these factors and account for potential inconsistencies in post-consumer stream plastics. Additive methods such as integrating nanocomposites into the asphalt mix should be highlighted in future works, while being considerate to limit the contribution of increased chemical toxicity.

Increased attention to environmental impact should also be a key area of future work. The characterization of MP and other chemical leachates yields contrasting results, proposing a direction for future research. It is also imperative to establish more concrete environmental benchmarks, such as MP pollution thresholds, to calibrate environmental hazards in a standardized way. Lastly, sustainability research on plastic roads should emphasize developing long-term, circular pipelines from raw plastic waste to pavement rejuvenation cycles. This includes viably repurposing plastic-modified pavements with RAP technology, conducting further investigations on the poten-

tial mobilization of toxic compounds, and developing effective strategies to mitigate them.

Overall, the challenges of plastic road implementation are the chemical persistence and toxicity hazards of incorporating plastic into geological infrastructure. Future work should highlight these crucial factors. Moving towards a circular economy necessitates greater emphasis on upstream recycling efforts with downstream solutions as temporary alternatives. Plastic roads should be a part of a larger environmental strategy to curb the U.S. plastic crisis, treated as a temporary solution to alleviate the burden from landfills, while sustainable solutions in upcycling and extended producer responsibility are emphasized for the long-term.

■ Acknowledgments

I would like to sincerely acknowledge Mary Conlon, my 9th-grade English 1 teacher, for facilitating the in-class forum research project that inspired me to write this review. I would also like to appreciate my forum groupmates for sparking insightful discussions and new ideas. Finally, thank you to my family for their continuous support and encouragement while I was writing my final manuscript.

■ References

1. Science Museum. The Age of Plastic: from Parkesine to Pollution. Science Museum. <https://www.sciencemuseum.org.uk/objects-and-stories/chemistry/age-plastic-parkesine-pollution>.
2. U.S. Environmental Protection Agency. Plastics: Material-Specific Data. Environmental Protection Agency. <https://www.epa.gov/facts-and-figures-about-materials-waste-and-recycling/plastics-material-specific-data> (accessed 2025-07-13).
3. Wakefield, F. World Recycling Facts for 2022: Plastic, Paper and More. World Economic Forum. <https://www.weforum.org/stories/2022/06/recycling-global-statistics-facts-plastic-paper/> (accessed 2025-07-13).
4. National Renewable Energy Laboratory. News Release: NREL Calculates Lost Value of Landfilled Plastic in the U.S. National Renewable Energy Laboratory. <https://www.nrel.gov/news/detail/press/2022/nrel-calculates-lost-value-of-landfilled-plastic-in-us> (accessed 2025-06-29).
5. OECD. Plastic Pollution Is Growing Relentlessly as Waste Management and Recycling Fall Short, Says OECD. OECD. <https://www.oecd.org/en/about/news/press-releases/2022/02/plastic-pollution-is-growing-relentlessly-as-waste-management-and-recycling-fall-short.html> (accessed 2025-06-29).
6. Nielsen, T. D.; Hasselbalch, J.; Holmberg, K.; Stripple, J. Politics and the Plastic Crisis: A Review throughout the Plastic Life Cycle. *Wiley Interdisciplinary Reviews: Energy and Environment* **2019**, 9 (1). <https://doi.org/10.1002/wene.360>.
7. MacLeod, M.; Arp, H. P. H.; Tekman, M. B.; Jahnke, A. The Global Threat from Plastic Pollution. *Science* **2021**, 373 (6550), 61–65. <https://doi.org/10.1126/science.abg5433>.
8. Ragusa, A.; Svelato, A.; Santacroce, C.; Catalano, P.; Notarstefano, V.; Carnevali, O.; Papa, F.; Rongioletti, M. C. A.; Baiocco, F.; Draghi, S.; D'Amore, E.; Rinaldo, D.; Matta, M.; Giorgini, E. Plasticenta: First Evidence of Microplastics in Human Placenta. *Environment International* **2021**, 146 (106274), 106274. <https://doi.org/10.1016/j.envint.2020.106274>.
9. Prata, J. C.; da Costa, J. P.; Lopes, I.; Duarte, A. C.; Rocha-Santos, T. Environmental Exposure to Microplastics: An Overview on Possi-

- ble Human Health Effects. *Science of the Total Environment* **2020**, 702 (134455). <https://doi.org/10.1016/j.scitotenv.2019.134455>.
10. Pauly, J. L.; Stegmeier, S. J.; Allaart, H. A.; Cheney, R. T.; Zhang, P. J.; Mayer, A. G.; Streck, R. J. Inhaled Cellulosic and Plastic Fibers Found in Human Lung Tissue. *Cancer Epidemiology, Biomarkers & Prevention* **1998**, 7 (5), 419–428.
 11. Alves, B. Topic: Plastic Waste in the United States. Statista. <https://www.statista.com/topics/5127/plastic-waste-in-the-unit-ed-states/> (accessed 2025-07-29).
 12. Illinois Environmental Protection Agency. 2023 Landfill Capacity Report. Illinois Environmental Protection Agency. <https://epa.illinois.gov/topics/waste-management/landfills/landfill-capacity/2023.html> (accessed 2025-07-13).
 13. Ágnes, N.; Rajmund, K. The Environmental Impact of Plastic Waste Incineration. *AARMS – Academic and Applied Research in Military and Public Management Science* **2016**, 15 (3), 231–237231–237. <https://doi.org/10.32565/aarms.2016.3.3>.
 14. Movilla-Quesada, D.; Lagos-Varas, M.; Raposeiras, A. C.; Muñoz-Cáceres, O.; Andrés-Valeri, V. C.; Aguilar-Vidal, C. Analysis of Greenhouse Gas Emissions and the Environmental Impact of the Production of Asphalt Mixes Modified with Recycled Materials. *Sustainability* **2021**, 13 (14), 8081. <https://doi.org/10.3390/su13148081>.
 15. Xu, F.; Zhao, Y.; Li, K. Using Waste Plastics as Asphalt Modifier: A Review. *Materials* **2021**, 15 (1), 110. <https://doi.org/10.3390/ma15010110>.
 16. California Department of Transportation. Caltrans Repaves Roadway with Recycled Plastic Bottles. California Department of Transportation. <https://dot.ca.gov/news-releases/news-release-2020-024> (accessed 2025-07-21).
 17. Corbley, A. 5 U.S. States Are Repaving Roads with Unrecyclable Plastic Waste—And Results Are Impressive. Good News Network. <https://www.goodnewsnetwork.org/these-5-u-s-states-are-repaving-roads-this-year-with-unrecyclable-plastic-waste-the-results-are-impressive/> (accessed 2025-07-21).
 18. Hawaii Department of Transportation. HDOT Testing Asphalt Mixes Designed to Improve Pavement and the Environment. Hawaii Department of Transportation. <https://hidot.hawaii.gov/administration/hdot-testing-asphalt-mixes-designed-to-improve-pavement-and-the-environment/> (accessed 2025-07-21).
 19. U.S. Chamber of Commerce Foundation. Dow: Using Recycled Plastics to Pave Roads and Parking Lots in Michigan. U.S. Chamber of Commerce Foundation. <https://www.uschamberfoundation.org/disasters/dow-using-recycled-plastics-pave-roads-and-parking-lots-michigan> (accessed 2025-07-14).
 20. Abd Karim, S. B.; Norman, S.; Koting, S.; Simarani, K.; Loo, S.-C.; Mohd Rahim, F. A.; Ibrahim, M. R.; Md Yusoff, N. I.; Nagor Mohamed, A. H. Plastic Roads in Asia: Current Implementations and Should It Be Considered? *Materials* **2023**, 16 (16), 5515. <https://doi.org/10.3390/ma16165515>.
 21. Buhl, Z. MnDOT finds PMBs Extend Road Service Life by Six Years – Asphalt Materials, Inc. Asphalt Materials, Inc. <https://asphalt-materials.com/mndot-finds-pmb-extend-road-service-life/> (accessed 2025-06-30).
 22. Mihai, M.; Tianhao, Y.; Mugurel, T.; Manik, B.; Vishruthi, M.; Manik, C. Cost Estimate of B vs. C Grade Asphalt Binders | MnDOT Digital Library. Mndot.gov. <https://mdl.mndot.gov/items/202319> (accessed 2025-06-30).
 23. Eklöf, K.; Wendel, M. Lifetime Performance of Polymer-Modified Bitumen in Swedish Roads Evaluated with Survival Analysis. *Transportation Research Record Journal of the Transportation Research Board* **2023**, 2678 (4), 106–117. <https://doi.org/10.1177/03611981231183713>.
 24. Bae, S.-H.; Chae, E.; Park, Y.-S.; Lee, S.-W.; Yun, J.-H.; Choi, S.-S. Characteristics of Tire-Road Wear Particles (TRWPs) and Road Pavement Wear Particles (RPWPs) Generated through a Novel Tire Abrasion Simulator Based on Real Road Pavement Conditions. *Science of the Total Environment* **2024**, 944, 173948–173948. <https://doi.org/10.1016/j.scitotenv.2024.173948>.
 25. Rødland, E. S.; Gustafsson, M.; Jaramillo-Vogel, D.; Järskog, I.; Müller, K.; Rauer, C.; Rausch, J.; Wagner, S. Analytical Challenges and Possibilities for the Quantification of Tire-Road Wear Particles. *TrAC Trends in Analytical Chemistry* **2023**, 165, 117121–117121. <https://doi.org/10.1016/j.trac.2023.117121>.
 26. Rødland, E. S.; Christian Lind, O.; Reid, M. J.; Sørli Heier, L.; Skogsberg, E.; Snilsberg, B.; Gryteselv, D.; Meland, S. Characterization of Tire and Road Wear Microplastic Particle Contamination in a Road Tunnel: From Surface to Release. *Journal of Hazardous Materials* **2022**, 435, 129032–129032. <https://doi.org/10.1016/j.jhazmat.2022.129032>.
 27. Ren, S.-Y.; Sun, Q.; Ni, H.-G.; Wang, J. A Minimalist Approach to Quantify Emission Factor of Microplastic by Mechanical Abrasion. *Chemosphere* **2020**, 245, 125630. <https://doi.org/10.1016/j.chemosphere.2019.125630>.
 28. Enfrin, M.; Myszk, R.; Giustozzi, F. Paving Roads with Recycled Plastics: Microplastic Pollution or Eco-Friendly Solution? *Journal of Hazardous Materials* **2022**, 437, 129334. <https://doi.org/10.1016/j.jhazmat.2022.129334>.
 29. Rincón-Esteva, J. A.; González-Salcedo, E. V.; Rondón-Quintana, H. A.; Reyes-Lizcano, F. A.; Bastidas-Martínez, J. G. Mechanical Behavior of Low-Density Polyethylene Waste Modified Hot Mix Asphalt. *Sustainability* **2022**, 14 (7), 4229. <https://doi.org/10.3390/su14074229>.
 30. Meng, Y.; Lai, J.; Mo, S.; Fang, G.; Deng, S.; Wei, X.; Yang, F. Investigating the Deterioration Mechanism of Adhesion between Asphalt and Aggregate Interface under Acid Rain Erosion. *Applied Surface Science* **2023**, 639, 158171. <https://doi.org/10.1016/j.apsusc.2023.158171>.
 31. Zou, Y.; Amirkhanian, S.; Xu, S.; Li, Y.; Wang, Y.; Zhang, J. Effect of Different Aqueous Solutions on Physicochemical Properties of Asphalt Binder. *Construction and Building Materials* **2021**, 286, 122810–122810. <https://doi.org/10.1016/j.conbuildmat.2021.122810>.
 32. Li, Y.; Madanat, S. A Steady-State Solution for the Optimal Pavement Resurfacing Problem. *Transportation Research Part A: Policy and Practice* **2002**, 36 (6), 525–535. [https://doi.org/10.1016/s0965-8564\(01\)00020-9](https://doi.org/10.1016/s0965-8564(01)00020-9).
 33. Jia Boom, Y.; Xuan Lu, D.; Enfrin, M.; Swaney, M.; Masood, H.; Kumar Pramanik, B.; Robert, D.; Giustozzi, F. Engineering Properties, Microplastics and Emissions Assessment of Recycled Plastic Modified Asphalt Mixtures. *Science of The Total Environment* **2023**, 893, 164869–164869. <https://doi.org/10.1016/j.scitotenv.2023.164869>.
 34. Harvey, J. T.; Tsai, B.-W. Effects of Asphalt Content and Air Void Content on Mix Fatigue and Stiffness. *Transportation Research Record: Journal of the Transportation Research Board* **1996**, 1543 (1), 38–45. <https://doi.org/10.1177/0361198196154300105>.
 35. Budziński, B.; Ratajczak, M.; Majer, S.; Wilmański, A. Influence of Bitumen Grade and Air Voids on Low-Temperature Cracking of Asphalt. *Case Studies in Construction Materials* **2023**, 19, e02255. <https://doi.org/10.1016/j.cscm.2023.e02255>.
 36. Arambula, E.; Masad, E.; Martin, A. E. Influence of Air Void Distribution on the Moisture Susceptibility of Asphalt Mixes. *Journal of Materials in Civil Engineering* **2007**, 19 (8), 655–664. [https://doi.org/10.1061/\(asce\)0899-1561\(2007\)19:8\(655\)](https://doi.org/10.1061/(asce)0899-1561(2007)19:8(655)).

37. Zaltuom, A. M. A Review Study of the Effect of Air Voids on Asphalt Pavement Life. *AJRR Proceedings*. <https://books.ajrr.org/index.php/press/catalog/book/34/chapter/194> (accessed 2025-07-20).
38. Duan, Y.; Wu, K.; Serrat, C.; Fabricio Arteaga-Larios; Brown, H.; DuBois, C. J.; Buttlar, W. G.; Deng, B. Assessment of Microplastics Production from Waste Plastics-Modified Asphalt Pavement. *Resources, Conservation and Recycling* **2024**, *202*, 107329–107329. <https://doi.org/10.1016/j.resconrec.2023.107329>.
39. Cui, P.; Schito, G.; Cui, Q. VOC Emissions from Asphalt Pavement and Health Risks to Construction Workers. *Journal of Cleaner Production* **2020**, *244*, 118757. <https://doi.org/10.1016/j.jclepro.2019.118757>.
40. Collins, J. F.; Brown, J. P.; Dawson, S. V.; Marty, M. A. Risk Assessment for Benzo[A]Pyrene. *Regulatory Toxicology and Pharmacology* **1991**, *13*(2), 170–184. [https://doi.org/10.1016/0273-2300\(91\)90020-v](https://doi.org/10.1016/0273-2300(91)90020-v).
41. Denissenko, M. F.; Pao, A.; Tang, M.; Pfeifer, G. P. Preferential Formation of Benzo[A]Pyrene Adducts at Lung Cancer Mutational Hotspots in P53. *Science* **1996**, *274* (5286), 430–432. <https://doi.org/10.1126/science.274.5286.430>.
42. Jia Boom, Y.; Enfrin, M.; Grist, S.; Giustozzi, F. Analysis of Possible Carcinogenic Compounds in Recycled Plastic Modified Asphalt. *Science of The Total Environment* **2023**, *858*, 159910–159910. <https://doi.org/10.1016/j.scitotenv.2022.159910>.
43. Carmona, E.; Rojo-Nieto, E.; Rummel, C. D.; Krauss, M.; Syberg, K.; Ramos, T. M.; Brosche, S.; Backhaus, T.; Almroth, B. C. A Dataset of Organic Pollutants Identified and Quantified in Recycled Polyethylene Pellets. *Data in Brief* **2023**, *51*, 109740. <https://doi.org/10.1016/j.dib.2023.109740>.
44. Tanushree Parsai; Dr. Arun Kumar. Human Risk Assessment: Toxicity Issues and Challenges Associated with Mixture of Chemicals Released during Plastic Reuse and Recycling; **2016**. <https://doi.org/10.3390/ecws-1-c003>.
45. Horodytska, O.; Cabanes, A.; Fullana, A. Non-Intentionally Added Substances (NIAS) in Recycled Plastics. *Chemosphere* **2020**, *251*, 126373. <https://doi.org/10.1016/j.chemosphere.2020.126373>.
46. Gazzotti, S.; De Felice, B.; Ortenzi, M. A.; Parolini, M. Approaches for Management and Valorization of Non-Homogeneous, Non-Recyclable Plastic Waste. *International Journal of Environmental Research and Public Health* **2022**, *19* (16), 10088. <https://doi.org/10.3390/ijerph191610088>.
47. Legret, M.; Odie, L.; Demare, D.; Jullien, A. Leaching of Heavy Metals and Polycyclic Aromatic Hydrocarbons from Reclaimed Asphalt Pavement. *Water Research* **2005**, *39* (15), 3675–3685. <https://doi.org/10.1016/j.watres.2005.06.017>.
48. Florida Pavement Preservation Council. FPPC Specs – The National Center for Pavement Preservation. Florida Pavement Preservation Council. <https://www.pavementpreservation.org/florida-pavement-preservation-council/fppc-specs/> (accessed 2025-07-02).
49. Zahoor, M.; Nizamuddin, S.; Madapusi, S.; Giustozzi, F. Sustainable Asphalt Rejuvenation Using Waste Cooking Oil: A Comprehensive Review. *Journal of Cleaner Production* **2021**, *278*, 123304. <https://doi.org/10.1016/j.jclepro.2020.123304>.
50. Di Mino, G.; Vijayan, V.; Eskandarsefat, S.; Venturini, L.; Mantalovas, K. Investigating the Multi-Recyclability of Recycled Plastic-Modified Asphalt Mixtures. *Infrastructures* **2023**, *8* (5), 84–84. <https://doi.org/10.3390/infrastructures8050084>.
51. Hoy, M.; Horpibulsuk, S.; Rachan, R.; Chinkulkijniwat, A.; Arulrajah, A. Recycled Asphalt Pavement – Fly Ash Geopolymers as a Sustainable Pavement Base Material: Strength and Toxic Leaching Investigations. *Science of The Total Environment* **2016**, *573*, 19–26. <https://doi.org/10.1016/j.scitotenv.2016.08.078>.
52. Lim, S. M.; He, M.; Hao, G.; Ng, A.; Ong, G. P. Recyclability Potential of Waste Plastic-Modified Asphalt Concrete with Consideration to Its Environmental Impact. *Construction & Building Materials* **2024**, *439*, 137299–137299. <https://doi.org/10.1016/j.conbuildmat.2024.137299>.
53. Spreadbury, C. J.; Clavier, K. A.; Lin, A. M.; Townsend, T. G. A Critical Analysis of Leaching and Environmental Risk Assessment for Reclaimed Asphalt Pavement Management. *Science of the Total Environment* **2021**, *775*, 145741. <https://doi.org/10.1016/j.scitotenv.2021.145741>.
54. Freire, A. C.; Neves, J. M. C. das; Roque, A. J.; Martins, I. M.; Lurdes Antunes, M. de. Feasibility Study of Milled and Crushed Reclaimed Asphalt Pavement for Application in Unbound Granular Layers. *Road Materials and Pavement Design* **2019**, *22* (7), 1500–1520. <https://doi.org/10.1080/14680629.2019.1701539>.
55. Muñoz, M.; Haag, R.; Figi, R.; Schreiner, C.; Zaubanis, M.; Cavalli, M. C.; Poulikakos, L. D.; Heeb, N. V. Environmental Impact of Rejuvenators in Asphalt Mixtures Containing High Reclaimed Asphalt Content. *Road Materials and Pavement Design* **2021**, *23* (6), 1400–1414. <https://doi.org/10.1080/14680629.2021.1891129>.
56. Kang, D.-H.; Gupta, S. C.; Bloom, P. R.; Ranaivoson, A. Z.; Roberson, R.; Siekmeier, J. Recycled Materials as Substitutes for Virgin Aggregates in Road Construction: II. Inorganic Contaminant Leaching. *Soil Science Society of America Journal* **2011**, *75* (4), 1276–1284. <https://doi.org/10.2136/sssaj2010.0296>.
57. Brantley, A. S.; Townsend, T. G. Leaching of Pollutants from Reclaimed Asphalt Pavement. *Environmental Engineering Science* **1999**, *16* (2), 105–116. <https://doi.org/10.1089/ees.1999.16.105>.
58. Mijic, Z.; Dayioglu, A. Y.; Hatipoglu, M.; Aydilek, A. H. Hydraulic and Environmental Impacts of Using Recycled Asphalt Pavement on Highway Shoulders. *Construction and Building Materials* **2020**, *234*, 117226. <https://doi.org/10.1016/j.conbuildmat.2019.117226>.
59. Saadeh, S.; Katawal, P.; Kabir, S. F.; Fini, E. H.; Titi, H. H. Performance of Reclaimed Asphalt Pavements Containing Recycled Waste Plastics. *Journal of Cleaner Production* **2024**, *442*, 140935–140935. <https://doi.org/10.1016/j.jclepro.2024.140935>.
60. Leng, Z.; Sreeram, A.; Padhan, R. K.; Tan, Z. Value-Added Application of Waste PET Based Additives in Bituminous Mixtures Containing High Percentage of Reclaimed Asphalt Pavement (RAP). *Journal of Cleaner Production* **2018**, *196*, 615–625. <https://doi.org/10.1016/j.jclepro.2018.06.119>.
61. Azadgoleh, M. A.; Mohammadi, M. M.; Ghodrati, A.; Sharifi, S. S.; Palizban, S. M. M.; Ahmadi, A.; Vahidi, E.; Ayar, P. Characterization of Contaminant Leaching from Asphalt Pavements: A Critical Review of Measurement Methods, Reclaimed Asphalt Pavement, Porous Asphalt, and Waste-Modified Asphalt Mixtures. *Water Research* **2022**, *219*, 118584. <https://doi.org/10.1016/j.watres.2022.118584>.
62. Santos, J.; Pham, A.; Stasinopoulos, P.; Giustozzi, F. Recycling Waste Plastics in Roads: A Life-Cycle Assessment Study Using Primary Data. *Science of The Total Environment* **2021**, *751*, 141842. <https://doi.org/10.1016/j.scitotenv.2020.141842>.
63. Yildirim, Y. Polymer Modified Asphalt Binders. *Construction and Building Materials* **2007**, *21* (1), 66–72. <https://doi.org/10.1016/j.conbuildmat.2005.07.007>.
64. Zhu, J.; Birgisson, B.; Kringos, N. Polymer Modification of Bitumen: Advances and Challenges. *European Polymer Journal* **2014**, *54*, 18–38. <https://doi.org/10.1016/j.eurpolymj.2014.02.005>.
65. Shivokhin, M.; García-Morales, M.; Partal, P.; Cuadri, A. A.; Gallegos, C. Rheological Behaviour of Polymer-Modified Bituminous Mastics: A Comparative Analysis between Physical and Chemical

- Modification. *Construction and Building Materials* **2012**, 27 (1), 234–240. <https://doi.org/10.1016/j.conbuildmat.2011.07.055>.
66. Revelli, V.; Kabir, S. F.; Ali, A.; Mehta, Y.; Cox, B. C. Understanding the Storage Stability of Polyethylene Modified Binders: A Laboratory Case Study Using Waste Plastics. *Journal of Materials in Civil Engineering* **2024**, 36 (4). <https://doi.org/10.1061/jmcee7.mteng-17334>.
 67. Haider, S.; Hafeez, I.; Jamal, M.; Ullah, R. Sustainable Use of Waste Plastic Modifiers to Strengthen the Adhesion Properties of Asphalt Mixtures. *Construction and Building Materials* **2020**, 235, 117496. <https://doi.org/10.1016/j.conbuildmat.2019.117496>.
 68. Casey, D.; McNally, C.; Gibney, A.; Gilchrist, M. D. Development of a Recycled Polymer Modified Binder for Use in Stone Mastic Asphalt. *Resources, Conservation and Recycling* **2008**, 52 (10), 1167–1174. <https://doi.org/10.1016/j.resconrec.2008.06.002>.
 69. Navarro, F.; Partal, P.; García-Morales, M.; Martín-Alfonso, M. J.; Martínez-Boza, F. J.; Gallegos, C.; Bordado, J. C.; Diogo, A. C. Bitumen Modification with Reactive and Non-Reactive (Virgin and Recycled) Polymers: A Comparative Analysis. *Journal of Industrial and Engineering Chemistry* **2009**, 15 (4), 458–464. <https://doi.org/10.1016/j.jiec.2009.01.003>.
 70. Stroup-Gardiner, M.; Brown, E. R. Segregation in Hot-mix Asphalt Pavements. Google Books. https://books.google.com/books?hl=en&lr=&id=0YNk9wAhojMC&oi=fnd&pg=PA82&dq=asphalt+segregation&ots=3tT8e07PMF&sig=PURZQ9Z-viQ9xXOV0YCPpH05_Jfk#v=onepage&q=asphalt%20segregation&cf=false (accessed 2025-08-05).
 71. Li, X.; Zhou, Z.; Lv, X.; Xiong, K.; Xinji, W.; You, Z. Temperature Segregation of Warm Mix Asphalt Pavement: Laboratory and Field Evaluations. *Construction and Building Materials* **2017**, 136, 436–445. <https://doi.org/10.1016/j.conbuildmat.2016.12.195>.
 72. Cong, L.; Yang, F.; Guo, G.; Ren, M.; Shi, J.; Tan, L. The Use of Polyurethane for Asphalt Pavement Engineering Applications: A State-of-the-Art Review. *Construction and Building Materials* **2019**, 225, 1012–1025. <https://doi.org/10.1016/j.conbuildmat.2019.07.213>.
 73. Mondal, P.; Khakhar, D. V. Hydraulic Resistance of Rigid Polyurethane Foams. III. Effect of Variation of the Concentration of Catalysts on Foam Structure and Properties. *Journal of Applied Polymer Science* **2004**, 93 (6), 2838–2843. <https://doi.org/10.1002/app.20763>.
 74. Park, S.-J.; Seo, M.-K. Chapter 6 – Element and Processing. ScienceDirect. <https://www.sciencedirect.com/science/article/abs/pii/B9780123750495000062> (accessed 2025-07-20).
 75. Wu, S.; Montalvo, L. Repurposing Waste Plastics into Cleaner Asphalt Pavement Materials: A Critical Literature Review. *Journal of Cleaner Production* **2021**, 280, 124355. <https://doi.org/10.1016/j.jclepro.2020.124355>.
 76. Vargas, M. A.; Vargas, M. A.; Sánchez-Sólis, A.; Manero, O. Asphalt/Polyethylene Blends: Rheological Properties, Microstructure and Viscosity Modeling. *Construction and Building Materials* **2013**, 45, 243–250. <https://doi.org/10.1016/j.conbuildmat.2013.03.064>.
 77. Qadir, A.; Imam, M. Use of Recycled Plastic Waste Aggregate as a Partial Substitution Material in Pavement Structure. In 2005 International Symposium on Pavement Recycling. Transportation Research Board: Washington, D.C., **2025**.
 78. Zoorob, S. E.; Suparma, L. B. Laboratory Design and Investigation of the Properties of Continuously Graded Asphaltic Concrete Containing Recycled Plastics Aggregate Replacement (Plastiphalt). *Cement and Concrete Composites* **2000**, 22 (4), 233–242. [https://doi.org/10.1016/S0958-9465\(00\)00026-3](https://doi.org/10.1016/S0958-9465(00)00026-3).
 79. Tayfur, S.; Ozen, H.; Aksoy, A. Investigation of Rutting Performance of Asphalt Mixtures Containing Polymer Modifiers. *Construction and Building Materials* **2007**, 21 (2), 328–337. <https://doi.org/10.1016/j.conbuildmat.2005.08.014>.
 80. Yan, K.; Xu, H.; You, L. Rheological Properties of Asphalts Modified by Waste Tire Rubber and Reclaimed Low Density Polyethylene. *Construction and Building Materials* **2015**, 83, 143–149. <https://doi.org/10.1016/j.conbuildmat.2015.02.092>.
 81. Al-Hadidy, A. I.; Yi-qiu, T. Effect of Polyethylene on Life of Flexible Pavements. *Construction and Building Materials* **2009**, 23 (3), 1456–1464. <https://doi.org/10.1016/j.conbuildmat.2008.07.004>.
 82. Hao, J.; Weiss, R. A. Viscoelastic and Mechanical Behavior of Hydrophobically Modified Hydrogels. *Macromolecules* **2011**, 44 (23), 9390–9398. <https://doi.org/10.1021/ma202130u>.
 83. García-Morales, M.; Partal, P.; Navarro, F.; Martínez-Boza, F. J.; Gallegos, C. Linear Viscoelasticity of Recycled EVA-Modified Bitumens. *Energy & Fuels* **2003**, 18 (2), 357–364. <https://doi.org/10.1021/ef034032u>.
 84. Ma, Y.; Zhou, H.; Jiang, X.; Polaczyk, P.; Xiao, R.; Zhang, M.; Huang, B. The Utilization of Waste Plastics in Asphalt Pavements: A Review. *Cleaner Materials* **2021**, 2, 100031. <https://doi.org/10.1016/j.clema.2021.100031>.
 85. Roja, K. L.; Rehman, A.; Ouederni, M.; Krishnamoorthy, S. K.; Abdala, A.; Masad, E. Influence of Polymer Structure and Amount on Microstructure and Properties of Polyethylene-Modified Asphalt Binders. *Materials and Structures* **2021**, 54 (2). <https://doi.org/10.1617/s11527-021-01683-0>.
 86. Mashaan, N.; Chegenizadeh, A.; Nikraz, H. A Comparison on Physical and Rheological Properties of Three Different Waste Plastic-Modified Bitumen. *Recycling* **2022**, 7 (2), 18. <https://doi.org/10.3390/recycling7020018>.
 87. Punith, V. S.; Veeraragavan, A. Evaluation of Reclaimed Polyethylene-Modified Asphalt Pavements. *Journal of Testing and Evaluation* **2010**, 38 (5), 102418. <https://doi.org/10.1520/jte102418>.
 88. Dalhat, M. A.; Al-Abdul Wahhab, H. I. Performance of Recycled Plastic Waste Modified Asphalt Binder in Saudi Arabia. *International Journal of Pavement Engineering* **2015**, 18 (4), 349–357. <https://doi.org/10.1080/10298436.2015.1088150>.
 89. Ho, S.; Church, R.; Klassen, K.; Law, B.; MacLeod, D.; Zanzotto, L. Study of Recycled Polyethylene Materials as Asphalt Modifiers. *Canadian Journal of Civil Engineering* **2006**, 33 (8), 968–981. <https://doi.org/10.1139/106-044>.
 90. Munaro, M.; Akcelrud, L. Correlations between Composition and Crystallinity of LDPE/HDPE Blends. *Journal of Polymer Research* **2007**, 15 (1), 83–88. <https://doi.org/10.1007/s10965-007-9146-2>.
 91. Cuadri, A. A.; Roman, C.; García-Morales, M.; Guisado, F.; Moreno, E.; Partal, P. Formulation and Processing of Recycled-Low-Density-Polyethylene-Modified Bitumen Emulsions for Reduced-Temperature Asphalt Technologies. *Chemical Engineering Science* **2016**, 156, 197–205. <https://doi.org/10.1016/j.ces.2016.09.018>.
 92. Naskar, M.; Chaki, T. K.; Reddy, K. S. Effect of Waste Plastic as Modifier on Thermal Stability and Degradation Kinetics of Bitumen/Waste Plastics Blend. *Thermochimica Acta* **2010**, 509 (1–2), 128–134. <https://doi.org/10.1016/j.tca.2010.06.013>.
 93. Padhan, R. K.; Sreeram, A. Enhancement of Storage Stability and Rheological Properties of Polyethylene (PE) Modified Asphalt Using Cross Linking and Reactive Polymer Based Additives. *Construction and Building Materials* **2018**, 188, 772–780. <https://doi.org/10.1016/j.conbuildmat.2018.08.155>.
 94. Araújo, J. R.; Waldman, W. R.; De Paoli, M. A. Thermal Properties of High Density Polyethylene Composites with Natural Fibres: Coupling Agent Effect. *Polymer Degradation and Stability*

- 2008, 93 (10), 1770–1775. <https://doi.org/10.1016/j.polymdegradstab.2008.07.021>.
95. Emblem, A. 13 – Plastics Properties for Packaging Materials. In ScienceDirect; Emblem, A., Emblem, H., Eds.; Woodhead Publishing: Cambridge, England, **2012**; pp. 287–309.
 96. White, G.; Hall, F. Laboratory Comparison of Wet-Mixing and Dry-Mixing of Recycled Waste Plastic for Binder and Asphalt Modification; Transportation Research International Documentation: Washington D.C., **2021**.
 97. Köfteci, S. Effect of HDPE Based Wastes on the Performance of Modified Asphalt Mixtures. *Procedia Engineering* **2016**, *161*, 1268–1274. <https://doi.org/10.1016/j.proeng.2016.08.567>.
 98. Punith, V. S.; Veeraragavan, A. Behavior of Reclaimed Polyethylene Modified Asphalt Cement for Paving Purposes. *Journal of Materials in Civil Engineering* **2010**, *23* (6), 833–845. [https://doi.org/10.1061/\(asce\)mt.1943-5533.0000235](https://doi.org/10.1061/(asce)mt.1943-5533.0000235).
 99. Cross, M. M. Viscosity, Molecular Weight and Chain Entanglement. *Polymer* **1970**, *11* (5), 238–244. [https://doi.org/10.1016/0032-3861\(70\)90034-0](https://doi.org/10.1016/0032-3861(70)90034-0).
 100. Appiah, J. K.; Berko-Boateng, V. N.; Tagbor, T. A. Use of Waste Plastic Materials for Road Construction in Ghana. *Case Studies in Construction Materials* **2017**, *6*, 1–7. <https://doi.org/10.1016/j.cscm.2016.11.001>.
 101. Hınıslıoğlu, S.; Açar, E. Use of Waste High Density Polyethylene as Bitumen Modifier in Asphalt Concrete Mix. *Materials Letters* **2004**, *58* (3–4), 267–271. [https://doi.org/10.1016/s0167-577x\(03\)00458-0](https://doi.org/10.1016/s0167-577x(03)00458-0).
 102. Melbouci, B.; Sadoun, S.; Bilek, A. Study of Strengthening of Recycled Asphalt Concrete by Plastic Aggregates. *International Journal of Pavement Research and Technology* **2014**, *7* (4), 280–286. [https://doi.org/10.6135/ijprt.org.tw/2014.7\(4\).280](https://doi.org/10.6135/ijprt.org.tw/2014.7(4).280).
 103. Costa, L. M. B.; Manuel, H. M. R.; Oliveira, J. R. M.; Fernandes, S. R. M. Incorporation of Waste Plastic in Asphalt Binders to Improve Their Performance in the Pavement. *International Journal of Pavement Research and Technology* **2013**, *6* (4), 457–464. [https://doi.org/10.6135/ijprt.org.tw/2013.6\(4\).457](https://doi.org/10.6135/ijprt.org.tw/2013.6(4).457).
 104. Gohil, S. V.; Suhail, S.; Rose, J.; Vella, T.; Nair, L. S. *Materials and Devices for Bone Disorders*, First.; Elsevier/Academic Press: Amsterdam, **2017**; pp. 349–403.
 105. Greene, J. P. *AUTOMOTIVE PLASTICS and COMPOSITES: Materials and Processing*, First.; Elsevier/William Andrew Publishing: S.L., **2021**; pp. 83–105.
 106. Ahmedzade, P.; Demirelli, K.; Günay, T.; Biryani, F.; Alqudah, O. Effects of Waste Polypropylene Additive on the Properties of Bituminous Binder. *Procedia Manufacturing* **2015**, *2*, 165–170. <https://doi.org/10.1016/j.promfg.2015.07.029>.
 107. Abtahi, S. M.; Esfandiarpour, S.; Kunt, M.; Hejazi, S. M.; Ebrahimi, M. G. Hybrid Reinforcement of Asphalt-Concrete Mixtures Using Glass and Polypropylene Fibers. *Journal of Engineered Fibers and Fabrics* **2013**, *8* (2), 155892501300800. <https://doi.org/10.1177/155892501300800203>.
 108. Buruiana, D. L.; Georgescu, P. L.; Carp, G. B.; Ghisman, V. Recycling Micro Polypropylene in Modified Hot Asphalt Mixture. *Scientific Reports* **2023**, *13* (1). <https://doi.org/10.1038/s41598-023-30857-9>.
 109. Tapkın, S. The Effect of Polypropylene Fibers on Asphalt Performance. *Building and Environment* **2008**, *43* (6), 1065–1071. <https://doi.org/10.1016/j.buildenv.2007.02.011>.
 110. Aschuri, I.; Yamin, A.; Widyasih, Y. D. The Use of Waste Plastic as a Partial Substitution Aggregate in Asphalt Concrete Pavement. *Jurnal Teknik Sipil* **2016**, *23* (1), 1–6. <https://doi.org/10.5614/jts.2016.23.1.1>.
 111. Otuoze, H.; Shuaibu, A. An Experimental Study on the Use of Polypropylene Waste in Bitumen Mix. *Nigerian Journal of Technology* **2017**, *36* (3), 677–685. <https://doi.org/10.4314/njt.v36i3.3>.
 112. M.A. Dalhat. Water Resistance and Characteristics of Asphalt Surfaces Treated with Micronized-Recycled-Polypropylene Waste: Super-Hydrophobicity. *Construction and Building Materials* **2021**, *285*, 122870–122870. <https://doi.org/10.1016/j.conbuildmat.2021.122870>.
 113. Chiang, T.-C.; Liu, H.-L.; Tsai, L.-C.; Jiang, T.; Ma, N.; Tsai, F.-C. Improvement of the Mechanical Property and Thermal Stability of Polypropylene/Recycled Rubber Composite by Chemical Modification and Physical Blending. *Scientific Reports* **2020**, *10* (1). <https://doi.org/10.1038/s41598-020-59191-0>.
 114. Vamegh, M.; Ameri, M.; Chavoshian Naeni, S. F. Experimental Investigation of Effect of PP/SBR Polymer Blends on the Moisture Resistance and Rutting Performance of Asphalt Mixtures. *Construction and Building Materials* **2020**, *253*, 119197. <https://doi.org/10.1016/j.conbuildmat.2020.119197>.
 115. Maine Department of Environmental Protection. Starting July 1, Enforcement of the Polystyrene Foam Ban begins, Maine DEP Media Release. Maine Department of Environmental Protection. <https://www.maine.gov/dep/news/news.html?id=5044853> (accessed 2025-07-20).
 116. The New York State Senate. TITLE 30 Expanded Polystyrene Foam Containers and Polystyrene Loose Fill Packaging Ban. The New York State Senate. <https://www.nysenate.gov/legislation/laws/ENV/A27T30> (accessed 2025-07-20).
 117. Kansas City News. “Trash to Roads” Pilot Program. Kansas City News. <https://www.kcmo.gov/Home/Components/News/News/2251/16?arch=1> (accessed 2025-07-21).
 118. Washington State Legislature. RCW 70A.245.070: Expanded Polystyrene Prohibitions—Penalty. Washington State Legislature. <https://app.leg.wa.gov/RCW/default.aspx?cite=70A.245.070> (accessed 2025-07-20).
 119. Maryland Department of the Environment. Polystyrene Food Service Products Ban. Maryland Department of the Environment. <https://mde.maryland.gov/programs/Land/RecyclingandOperationsprogram/Pages/Expanded-Polystyrene-Food-Service-Products-Ban.aspx> (accessed 2025-07-20).
 120. Pasztor, A. J.; Landes, B. G.; Karjala, P. J. Thermal Properties of Syndiotactic Polystyrene. *Thermochimica Acta* **1991**, *177*, 187–195. [https://doi.org/10.1016/0040-6031\(91\)80095-z](https://doi.org/10.1016/0040-6031(91)80095-z).
 121. Fang, C.; Jiao, L.; Hu, J.; Yu, Q.; Guo, D.; Zhou, X.; Yu, R. Viscoelasticity of Asphalt Modified with Packaging Waste Expanded Polystyrene. *Journal of Materials Science & Technology* **2014**, *30* (9), 939–943. <https://doi.org/10.1016/j.jmst.2014.07.016>.
 122. Mohd Hasan, M. R.; Colbert, B.; You, Z.; Jamshidi, A.; Heiden, P. A.; Hamzah, M. O. A Simple Treatment of Electronic-Waste Plastics to Produce Asphalt Binder Additives with Improved Properties. *Construction and Building Materials* **2016**, *110*, 79–88. <https://doi.org/10.1016/j.conbuildmat.2016.02.017>.
 123. Vila-Cortavarte, M.; Lastra-González, P.; Calzada-Pérez, M. Á.; Indacochea-Vega, I. Analysis of the Influence of Using Recycled Polystyrene as a Substitute for Bitumen in the Behaviour of Asphalt Concrete Mixtures. *Journal of Cleaner Production* **2018**, *170*, 1279–1287. <https://doi.org/10.1016/j.jclepro.2017.09.232>.
 124. Colbert, B. W.; You, Z. Properties of Modified Asphalt Binders Blended with Electronic Waste Powders. *Journal of Materials in Civil Engineering* **2012**, *24* (10), 1261–1267. [https://doi.org/10.1061/\(asce\)mt.1943-5533.0000504](https://doi.org/10.1061/(asce)mt.1943-5533.0000504).
 125. Lastra-González, P.; Calzada-Pérez, M. A.; Castro-Fresno, D.; Vega-Zamanillo, Á.; Indacochea-Vega, I. Comparative Analysis of the Performance of Asphalt Concretes Modified by Dry Way with

- Polymeric Waste. *Construction and Building Materials* **2016**, *112*, 1133–1140. <https://doi.org/10.1016/j.conbuildmat.2016.02.156>.
126. Vasudevan, R.; Ramalinga Chandra Sekar, A.; Sundarakannan, B.; Velkennedy, R. A Technique to Dispose Waste Plastics in an Ecofriendly Way – Application in Construction of Flexible Pavements. *Construction and Building Materials* **2012**, *28* (1), 311–320. <https://doi.org/10.1016/j.conbuildmat.2011.08.031>.
 127. Gutierrez-Velasquez, E. I.; Monteiro, S. N.; Colorado, H. A. Characterization of Expanded Polystyrene Waste as Binder and Coating Material. Case Studies in Construction Materials **2022**, *16*, e00804–e00804. <https://doi.org/10.1016/j.cscm.2021.e00804>.
 128. M. A. S. Maia, M.; Dinis-Almeida, M.; C. G. Martinho, F. The Influence of the Affinity between Aggregate and Bitumen on the Mechanical Performance Properties of Asphalt Mixtures. *Materials* **2021**, *14* (21), 6452–6452. <https://doi.org/10.3390/ma14216452>.
 129. Rochman, C. M.; Manzano, C.; Hentschel, B. T.; Simonich, S. L. M.; Hoh, E. Polystyrene Plastic: A Source and Sink for Polycyclic Aromatic Hydrocarbons in the Marine Environment. *Environmental Science & Technology* **2013**, *47* (24), 13976–13984. <https://doi.org/10.1021/es403605f>.
 130. Turner, A. Foamed Polystyrene in the Marine Environment: Sources, Additives, Transport, Behavior, and Impacts. *Environmental Science & Technology* **2020**, *54* (17), 10411–10420. <https://doi.org/10.1021/acs.est.0c03221>.
 131. Ghabchi, R.; Dharmarathna, C. P.; Mihandoust, M. Feasibility of Using Micronized Recycled Polyethylene Terephthalate (PET) as an Asphalt Binder Additive: A Laboratory Study. *Construction and Building Materials* **2021**, *292*, 123377. <https://doi.org/10.1016/j.conbuildmat.2021.123377>.
 132. Ahmadinia, E.; Zargar, M.; Karim, M. R.; Abdelaziz, M.; Ahmadinia, E. Performance Evaluation of Utilization of Waste Polyethylene Terephthalate (PET) in Stone Mastic Asphalt. *Construction and Building Materials* **2012**, *36*, 984–989. <https://doi.org/10.1016/j.conbuildmat.2012.06.015>.
 133. Nisticò, R. Polyethylene Terephthalate (PET) in the Packaging Industry. *Polymer Testing* **2020**, *90* (0142–9418), 106707. <https://doi.org/10.1016/j.polymertesting.2020.106707>.
 134. Soong, Y.-H. V.; Sobkowicz, M. J.; Xie, D. Recent Advances in Biological Recycling of Polyethylene Terephthalate (PET) Plastic Wastes. *Bioengineering* **2022**, *9* (3), 98. <https://doi.org/10.3390/bioengineering9030098>.
 135. U.S. Department of Transportation Federal Highway Administration. User Guidelines for Waste and Byproduct Materials in Pavement Construction. Federal Highway Administration Research and Technology. <https://www.fhwa.dot.gov/publications/research/infrastructure/structures/97148/st2.cfm> (accessed 2025-07-19).
 136. Gürü, M.; Çubuk, M. K.; Arslan, D.; Farzanian, S. A.; Bilici, İ. An Approach to the Usage of Polyethylene Terephthalate (PET) Waste as Roadway Pavement Material. *Journal of Hazardous Materials* **2014**, *279*, 302–310. <https://doi.org/10.1016/j.jhazmat.2014.07.018>.
 137. Sojobi, A. O.; Nwobodo, S. E.; Aladegboye, O. J. Recycling of Polyethylene Terephthalate (PET) Plastic Bottle Wastes in Bituminous Asphaltic Concrete. *Cogent Engineering* **2016**, *3* (1). <https://doi.org/10.1080/23311916.2015.1133480>.
 138. El-Naga, I. A.; Ragab, M. Benefits of Utilization the Recycle Polyethylene Terephthalate Waste Plastic Materials as a Modifier to Asphalt Mixtures. *Construction and Building Materials* **2019**, *219*, 81–90. <https://doi.org/10.1016/j.conbuildmat.2019.05.172>.
 139. Ameri, M.; Nasr, D. Performance Properties of Devulcanized Waste PET Modified Asphalt Mixtures. *Petroleum Science and Technology* **2016**, *35* (1), 99–104. <https://doi.org/10.1080/10916466.2016.1251457>.
 140. Xu, X.; Chen, G.; Wu, Q.; Leng, Z.; Chen, X.; Zhai, Y.; Tu, Y.; Peng, C. Chemical Upcycling of Waste PET into Sustainable Asphalt Pavement Containing Recycled Concrete Aggregates: Insight into Moisture-Induced Damage. *Construction and Building Materials* **2022**, *360*, 129632. <https://doi.org/10.1016/j.conbuildmat.2022.129632>.
 141. Choudhary, R.; Kumar, A.; Murkute, K. View of Properties of Waste Polyethylene Terephthalate (PET) Modified Asphalt Mixes: Dependence on PET Size, PET Content, and Mixing Process. *Periodica Polytechnica Civil Engineering* **2018**, *62* (3), 685–6933. <https://doi.org/10.3311/PPci.10797>.
 142. Mashaan, N. S.; Chegenizadeh, A.; Nikraz, H.; Rezagholilou, A. Investigating the Engineering Properties of Asphalt Binder Modified with Waste Plastic Polymer. *Ain Shams Engineering Journal* **2021**, *12* (2), 1569–1574. <https://doi.org/10.1016/j.asej.2020.08.035>.
 143. Moghaddam, T. B.; Karim, M. R.; Syammaun, T. Dynamic Properties of Stone Mastic Asphalt Mixtures Containing Waste Plastic Bottles. *Construction and Building Materials* **2012**, *34*, 236–242. <https://doi.org/10.1016/j.conbuildmat.2012.02.054>.
 144. Amobonye, A. E.; Bhagwat, P.; Singh, S.; Pillai, S. Biodegradability of Polyvinyl Chloride. *Biodegradability of Conventional Plastics* **2023**, 201–220. <https://doi.org/10.1016/b978-0-323-89858-4.00017-8>.
 145. Turner, A.; Filella, M. Polyvinyl Chloride in Consumer and Environmental Plastics, with a Particular Focus on Metal-Based Additives. *Environmental Science: Processes & Impacts* **2021**, 1376–1384. <https://doi.org/10.1039/d1em00213a>.
 146. Liu, P.; Zhu, L.; Fang, Y.; Zhang, H.; Chen, D.; Xu, K.; Chen, M. Hydroxylbenzylthioethers as Novel Organic Thermal Stabilizers for Rigid PVC. *Polymer Degradation and Stability* **2007**, *92* (3), 503–508. <https://doi.org/10.1016/j.polymdegradstab.2006.07.013>.
 147. Fakhri, M.; Shahryari, E.; Ahmadi, T. Investigate the Use of Recycled Polyvinyl Chloride (PVC) Particles in Improving the Mechanical Properties of Stone Mastic Asphalt (SMA). *Construction and Building Materials* **2022**, *326*, 126780. <https://doi.org/10.1016/j.conbuildmat.2022.126780>.
 148. Köfteci, S.; Ahmedzade, P.; Kultayev, B. Performance Evaluation of Bitumen Modified by Various Types of Waste Plastics. *Construction and Building Materials* **2014**, *73*, 592–602. <https://doi.org/10.1016/j.conbuildmat.2014.09.067>.
 149. Arabani, M.; Taleghani, M. Y. Rutting Behavior of Hot Mix Asphalt Modified by Polyvinyl Chloride Powder. *Petroleum Science and Technology* **2017**, *35* (15), 1621–1626. <https://doi.org/10.1080/10916466.2017.1336772>.
 150. Salman, N.; Jaleel, Z. Effects of Waste PVC Addition on the Properties of (40–50) Grade Asphalt. *MATEC Web of Conferences* **2018**, *162*, 01046. <https://doi.org/10.1051/mateconf/201816201046>.
 151. Mahmoud Abdel-Halim Abdel-Goad. The Effect of Soaking Time on the Adhesive Strength of Polyvinylchloride Plates. *International Journal of Engineering and Information Systems* **2022**, *6* (5), 35–38.
 152. Behl, A.; Sharma, G.; Kumar, G. A Sustainable Approach: Utilization of Waste PVC in Asphalt of Roads. *Construction and Building Materials* **2014**, *54*, 113–117. <https://doi.org/10.1016/j.conbuildmat.2013.12.050>.
 153. Fang, C.; Zhou, S.; Zhang, M.; Zhao, S. Modification of Waterproofing Asphalt by PVC Packaging Waste. *Journal of Vinyl and Additive Technology* **2009**, *15* (4), 229–233. <https://doi.org/10.1002/vnl.20204>.

154. Ziari, H.; Nasiri, E.; Amini, A.; Ferdosian, O. The Effect of EAF Dust and Waste PVC on Moisture Sensitivity, Rutting Resistance, and Fatigue Performance of Asphalt Binders and Mixtures. *Construction and Building Materials* **2019**, *203*, 188–200. <https://doi.org/10.1016/j.conbuildmat.2019.01.101>.
155. Xia, C.; Cong, B.; Lv, S.; Wang, D.; Liu, B.; Zhao, T.; Jiang, X. Characterization of Fatigue Damage in High-Modulus Asphalt Mixtures at Different Aging Degrees. *Construction and Building Materials* **2025**, *472*, 140961. <https://doi.org/10.1016/j.conbuildmat.2025.140961>.
156. Rasel, H. M.; Rahman, M. N.; Ahmed, T. U. Study of Effects of Waste PVC on the Properties of Bituminous Mixes. *SAMRID-DHI – a Journal of Physical Sciences, Engineering, and Technology* **2011**, *2* (02), 17–23. <https://doi.org/18090/samriddhi.v2i2.1601>.
157. Rahman, Md. N.; Ahmeduzzaman, M.; A. Sobhan, M.; U. Ahmed, T. Performance Evaluation of Waste Polyethylene and PVC on Hot Asphalt Mixtures. *American Journal of Civil Engineering and Architecture* **2013**, *1* (5), 97–102. <https://doi.org/10.12691/ajcea-1-5-2>.
158. Nizamuddin, S.; Boom, Y. J.; Giustozzi, F. Sustainable Polymers from Recycled Waste Plastics and Their Virgin Counterparts as Bitumen Modifiers: A Comprehensive Review. *Polymers* **2021**, *13* (19), 3242. <https://doi.org/10.3390/polym13193242>.
159. Chan, H. S. O. Measurement of Hydrochloric Acid Emission from Burning PVC Compounds. *Journal of Fire Sciences* **1984**, *2* (2), 106–122. <https://doi.org/10.1177/073490418400200203>.
160. Jiang, D.; Cao, Z.; Zhou, B.; Gong, G.; Huang, Z.; Wang, C. Environmental Impact Assessment of Waste Plastic Modified Asphalt Mixtures: VOCs Emission Characteristics, Environmental-Health Risks, and Underlying Mechanism. *Environmental Research* **2025**, *279*, 121642. <https://doi.org/10.1016/j.envres.2025.121642>.
161. Evans, C. D.; Monteith, D. T.; Fowler, D.; Cape, J. N.; Brayshaw, S. Hydrochloric Acid: An Overlooked Driver of Environmental Change. *Environmental Science & Technology* **2011**, *45* (5), 1887–1894. <https://doi.org/10.1021/es103574u>.
162. Ragab, A. A.; Mohammedy, M. M.; M. El-Shafie. Using Waste Flexible Polyvinyl Chloride Treated with DOP/Calcium Hydroxide for Enriching the Performance of Oxidizing Bitumen. *Journal of Thermal Analysis and Calorimetry* **2018**, *136* (3), 1079–1091. <https://doi.org/10.1007/s10973-018-7754-1>.
163. Meincke, O.; Kaempfer, D.; Weickmann, H.; Friedrich, C.; Vathauer, M.; Warth, H. Mechanical Properties and Electrical Conductivity of Carbon-Nanotube Filled Polyamide-6 and Its Blends with Acrylonitrile/Butadiene/Styrene. *Polymer* **2004**, *45* (3), 739–748. <https://doi.org/10.1016/j.polymer.2003.12.013>.
164. Tiganis, B. E.; Burn, L. S.; Davis, P.; Hill, A. J. Thermal Degradation of Acrylonitrile–Butadiene–Styrene (ABS) Blends. *Polymer Degradation and Stability* **2002**, *76* (3), 425–434. [https://doi.org/10.1016/s0141-3910\(02\)00045-9](https://doi.org/10.1016/s0141-3910(02)00045-9).
165. Yang, Q.; Lin, J.; Wang, X.; Wang, D.; Xie, N.; Shi, X. A Review of Polymer-Modified Asphalt Binder: Modification Mechanisms and Mechanical Properties. *Cleaner Materials* **2024**, *12*, 100255. <https://doi.org/10.1016/j.clema.2024.100255>.
166. Fonseca, M.; Capitão, S.; Almeida, A.; Picado-Santos, L. Influence of Plastic Waste on the Workability and Mechanical Behaviour of Asphalt Concrete. *Applied Sciences* **2022**, *12* (4), 2146. <https://doi.org/10.3390/app12042146>.
167. Singh, P. K.; Suman, S. K.; Kumar, M. Influence of Recycled Acrylonitrile Butadiene Styrene (ABS) on the Physical, Rheological and Mechanical Properties of Bitumen Binder. *Transportation Research Procedia* **2020**, *48*, 3668–3677. <https://doi.org/10.1016/j.trpro.2020.08.081>.
168. García-Morales, M.; Partal, P.; Navarro, F. J.; Gallegos, C. Effect of Waste Polymer Addition on the Rheology of Modified Bitumen. *Fuel* **2006**, *85* (7–8), 936–943. <https://doi.org/10.1016/j.fuel.2005.09.015>.
169. Audy, R.; Enfrin, M.; Boom, Y. J.; Giustozzi, F. Selection of Recycled Waste Plastic for Incorporation in Sustainable Asphalt Pavements: A Novel Multi-Criteria Screening Tool Based on 31 Sources of Plastic. *Science of The Total Environment* **2022**, *829*, 154604. <https://doi.org/10.1016/j.scitotenv.2022.154604>.
170. Stephens, B.; Azimi, P.; El Orch, Z.; Ramos, T. Ultrafine Particle Emissions from Desktop 3D Printers. *Atmospheric Environment* **2013**, *79*, 334–339. <https://doi.org/10.1016/j.atmosenv.2013.06.050>.
171. Unwin, J.; Coldwell, M. R.; Keen, C.; McAlinden, J. J. Airborne Emissions of Carcinogens and Respiratory Sensitizers during Thermal Processing of Plastics. *The Annals of Occupational Hygiene* **2012**, *57* (3), 399–406. <https://doi.org/10.1093/annhyg/mes078>.
172. Wojnowski, W.; Marć, M.; Kalinowska, K.; Kosmela, P.; Zabiegała, B. Emission Profiles of Volatiles during 3D Printing with ABS, ASA, Nylon, and PETG Polymer Filaments. *Molecules* **2022**, *27* (12), 3814–3814. <https://doi.org/10.3390/molecules27123814>.
173. Wang, W.; Wang, C. *The Design and Manufacture of Medical Devices*, 1st ed.; Woodhead Publishing: Oxford, **2012**; pp. 115–151.
174. Das, A.; Mahanwar, P. A Brief Discussion on Advances in Polyurethane Applications. *Advanced Industrial and Engineering Polymer Research* **2020**, *3* (3), 93–101. <https://doi.org/10.1016/j.aiepr.2020.07.002>.
175. Yu, X.; Chen, B.; Dong, F.; Zu, Y.; Wang, S.; Si, J.; Huang, J. Repurposing Waste Polyurethane into Cleaner Asphalt Pavement Materials: Laboratory Investigation. *Journal of Materials in Civil Engineering* **2023**, *35* (9). <https://doi.org/10.1061/jmcee7.mteng-15892>.
176. Wei, K.; Wang, X.; Ma, B.; Shi, W.; Duan, S.; Liu, F. Study on Rheological Properties and Phase-Change Temperature Control of Asphalt Modified by Polyurethane Solid–Solid Phase Change Material. *Solar Energy* **2019**, *194*, 893–902. <https://doi.org/10.1016/j.solener.2019.11.007>.
177. Sun, M.; Zheng, M.; Qu, G.; Yuan, K.; Bi, Y.; Wang, J. Performance of Polyurethane Modified Asphalt and Its Mixtures. *Construction and Building Materials* **2018**, *191*, 386–397. <https://doi.org/10.1016/j.conbuildmat.2018.10.025>.
178. Hao, H.; Chen, Z.; Cong, P.; Han, Z. Rheological, Chemical and Short-Term Aging Properties of Waste Polyurethane Particles Modified Asphalt Binder with or without SBS. *Construction and Building Materials* **2022**, *357*, 129363. <https://doi.org/10.1016/j.conbuildmat.2022.129363>.
179. U.S. Department of Transportation Federal Highway Administration. Asphalt Resources. U.S. Department of Transportation Federal Highway Administration. https://www.fhwa.dot.gov/pavement/tops/asphalt_resources.cfm (accessed 2025-08-16).
180. Yao, L.; Leng, Z.; Lan, J.; Chen, R.; Jiang, J. Environmental and Economic Assessment of Collective Recycling Waste Plastic and Reclaimed Asphalt Pavement into Pavement Construction: A Case Study in Hong Kong. *Journal of Cleaner Production* **2022**, *336*, 130405. <https://doi.org/10.1016/j.jclepro.2022.130405>.
181. Singh, A.; Gupta, A. Mechanical and Economical Feasibility of LDPE Waste-Modified Asphalt Mixtures: Pathway to Sustainable Road Construction. *Scientific Reports* **2024**, *14* (1). <https://doi.org/10.1038/s41598-024-75196-5>.
182. Forkenbrock, D. J.; Foster, N. S. J. Economic Benefits of a Corridor Highway Investment. *Transportation Research Part A: General* **1990**, *24* (4), 303–312. [https://doi.org/10.1016/0191-2607\(90\)90007-s](https://doi.org/10.1016/0191-2607(90)90007-s).

183. U.S. Department of Transportation Federal Highway Administration. Productivity and the Highway Network: A Look at the Economic Benefits to Industry from Investment in the Highway Network – Policy. Federal Highway Administration. <https://www.fhwa.dot.gov/policy/otps/060320b/> (accessed 2025-08-16).
184. Finnegan, D. J.; Forkenbrock, D. J.; Foster, N. S. J.; Pogue, T. F. Road Investment to Foster Local Economic Development. Iowa Research Online. https://iro.uiowa.edu/esploro/outputs/report/Road-Investment-to-Foster-Local-Economic/9983557196202771?institution=01IOWA_INST (accessed 2025-08-16).
185. Gertler, P. J.; Gonzalez-Navarro, M.; Gračner, T.; Rothenberg, A. D. Road Maintenance and Local Economic Development: Evidence from Indonesia's Highways. *Journal of Urban Economics* 2024, 143, 103687. <https://doi.org/10.1016/j.jue.2024.103687>.
186. U.S. Department of Transportation Federal Highway Administration. Economic Strength: Creating Reliable and Efficient Access to Resources, Markets and Good-Paying Jobs. U.S. Department of Transportation. <https://www.transportation.gov/economic-strength-creating-reliable-efficient> (accessed 2025-08-16).
187. Hooper, E.; Peters, S.; Pintus, P. A. The Impact of Infrastructure Investments on Income Inequality: Evidence from US States. *Economics of Transition and Institutional Change* 2020, 29 (2), 227–256. <https://doi.org/10.1111/ecot.12266>.
188. American Society of Civil Engineers. ASCE Report Card Gives U.S. Infrastructure Highest-Ever C Grade. American Society of Civil Engineers (ASCE). <https://www.asce.org/publications-and-news/civil-engineering-source/society-news/article/2025/03/25/asce-report-card-gives-us-infrastructure-highest-ever-c-grade> (accessed 2025-08-16).
189. U.S. Department of Transportation. Bipartisan Infrastructure Law (BIL) / Infrastructure Investment and Jobs Act (IIJA). U.S. Department of Transportation Pipeline and Hazardous Materials Safety Administration. <https://www.phmsa.dot.gov/legislative-mandates/bipartisan-infrastructure-law-bil-infrastructure-investment-and-jobs-act-iija> (accessed 2025-08-16).
190. Dow Chemical Company. Dow Completes Roads Improved with Recycled Plastic. Dow. <https://corporate.dow.com/en-us/news/press-releases/dow-completes-roads-improved-with-recycled-plastic.html> (accessed 2025-08-16).
191. MacRebur. Sustainable Polymers for Asphalt. MacRebur. <https://www.macrebur.com/> (accessed 2025-08-16).
192. House, T. W. FACT SHEET: Biden-Harris Administration Transforms Nation's Infrastructure, Celebrates Historic Progress in Rebuilding America for the Three-Year Anniversary of the Bipartisan Infrastructure Law. The White House. <https://bidenwhitehouse.archives.gov/briefing-room/statements-releases/2024/11/15/fact-sheet-biden-%E2%81%A0harris-administration-transforms-nations-infrastructure-celebrates-historic-progress-in-rebuilding-america-for-the-three-year-anniversary-of-the-bipartisan-infrac/> (accessed 2025-08-16).
193. U.S. Department of Transportation Federal Highway Administration. Trump's Transportation Secretary Sean P. Duffy Slashes Biden-Era Greenhouse Gas Rule. Federal Highway Administration. <https://highways.dot.gov/node/132596> (accessed 2025-08-16).
194. EHN Curators. Trump Administration Weakens Environmental Review Process, Creating Uncertainty. Environmental Health News. <https://www.ehn.org/trump-administration-weakens-environmental-review-process-creating-uncertainty> (accessed 2025-08-16).
195. Alms, N. Trump Administration Debuts Permitting Modernization Plan, Even as Staff Cuts Could Jeopardize It. Nextgov. <https://www.nextgov.com/modernization/2025/06/trump-administration-debuts-permitting-modernization-plan-even-staff-cuts-could-jeopardize-it/405848/> (accessed 2025-08-16).
196. American Association of State Highway and Transportation Officials (AASHTO). White House Unveils Infrastructure Permitting Reform Plan. AASHTO Journal. <https://aashtojournal.transportation.org/white-house-unveils-infrastructure-permitting-reform-plan/> (accessed 2025-08-16).
197. Conlon, K. Plastic Roads: Not All They're Paved up to Be. *International Journal of Sustainable Development & World Ecology* 2021, 29 (1), 80–83. <https://doi.org/10.1080/13504509.2021.1915406>.
198. Bergmann, M.; Peter, H.; Almroth, B. C.; Cowger, W.; Eriksen, M.; Dey, T.; Gündoğdu, S.; Helm, R. R.; Krieger, A.; Syberg, K.; Tekman, M. B.; Thompson, R. C.; Villarrubia-Gómez, P.; Warrier, A. K.; Farrelly, T. Moving from Symptom Management to Upstream Plastics Prevention: The Fallacy of Plastic Clean-up Technology. *One Earth* 2023, 6 (11), 1439–1442. <https://doi.org/10.1016/j.oneear.2023.10.022>.
199. McKeen, L. W. The Effect of Long Term Thermal Exposure on Plastics and Elastomers, Second.; Elsevier, 2021; pp. 65–82.
200. Werbowski, L. M.; Gilbreath, A. N.; Munno, K.; Zhu, X.; Grbic, J.; Wu, T.; Sutton, R.; Sedlak, M. D.; Deshpande, A. D.; Rochman, C. M. Urban Stormwater Runoff: A Major Pathway for Anthropogenic Particles, Black Rubbery Fragments, and Other Types of Microplastics to Urban Receiving Waters. *ACS ES&T Water* 2021, 1 (6), 1420–1428. <https://doi.org/10.1021/acsestwater.1c00017>.
201. Bullard, J. E.; Ockelford, A.; O'Brien, P.; McKenna Neuman, C. Preferential Transport of Microplastics by Wind. *Atmospheric Environment* 2021, 245, 118038. <https://doi.org/10.1016/j.atmosenv.2020.118038>.
202. Hahladakis, J. N.; Velis, C. A.; Weber, R.; Iacovidou, E.; Purnell, P. An Overview of Chemical Additives Present in Plastics: Migration, Release, Fate and Environmental Impact during Their Use, Disposal and Recycling. *Journal of Hazardous Materials* 2018, 344 (344), 179–199. <https://doi.org/10.1016/j.jhazmat.2017.10.014>.

■ Author

Jasmine Liu is an 11th grader and passionate about green chemistry and sustainable engineering. She is working at the intersection of edge learning and water quality to design scalable algorithms for smart water monitoring. She is eager to contribute to developing chemical engineering solutions to address challenges in environmental sustainability.

Experiences of Compassion Fatigue in Doctors: An Exploration of Triggers, Coping Strategies, and Culture

Lea C. Y. Yuen

The Perse School, Hills Road, Cambridge, Cambridgeshire, CB2 8QF, The UK; leayuency@gmail.com

Mentor: Sophie G. Dixon

ABSTRACT: There is currently a growing interest in exploring the effects of Compassion Fatigue (CF) in doctors due to its high prevalence and its association with declines in well-being amongst sufferers, as well as lower quality patient care. This study investigates how hospital doctors experience CF, what coping strategies they utilize, and how culture and medical training influence this. Semi-structured interviews were conducted with seven physicians who represented a range of specialties, years of experience, and cultural backgrounds. Data were analyzed using a hybrid thematic analysis approach. Five a priori (determined prior to analysis) themes and three posteriori (emerged from analysis) themes were identified and analyzed. The a priori themes included: Triggers of CF, Coping Strategies, Signs of CF, Influence of Cultural Background on CF, and Compassion as a Limited and Fluctuating Resource; the posteriori themes included: Lack of Preparation in Medical School, Attitudes About Mental Health, and Triggers of Compassion Satisfaction. Similar triggers and coping strategies were identified in the current study to those in the existing literature. However, the current study builds on this further by highlighting how experiences of CF differed across specialties, years of clinical experience, and cultural backgrounds. The findings suggest that support systems and intervention programs for CF among doctors should be tailored according to some of these factors, which should be an area of focus for future research.

KEYWORDS: Psychology, Health Psychology, Compassion Fatigue, Triggers, Coping Strategies, Culture.

■ Introduction

Doctors are exposed to suffering and trauma every day, which inevitably impacts their mental health. 62% of medical professionals claim to have experienced Compassion Fatigue (CF) as a result of their job.¹ According to Figley, Huggard, and Rees,² CF is defined as a reduction in the capacity or interest in bearing the suffering of others. Recent studies found that CF in medical professionals leads to stress, burnout, and a loss of compassion, resulting in poor quality care.^{3,4} CF is also found to be associated with high rates of staff turnover, potential financial implications for healthcare organizations, and concerns for patient safety and the well-being of doctors.⁴

The current study aims to investigate the personal insights of doctors working in public hospitals on the topic of CF. Although there is existing literature exploring the negative psychological, physical, and social effects of doctors experiencing CF, there is a dearth of research regarding the different ways physicians experience CF through in-depth self-report.^{5,6} While doctors frequently identify themselves as having experienced CF in survey studies,^{3,6,9,14,18,19} the way in which they experience these emotions and symptoms is often unexplored and is likely to vary significantly between individuals. For example, research has shown that doctors' perceptions of compassion are shaped by culture and framing, demonstrating that cultural background and training could profoundly impact how they experience CF.⁶ However, no research currently directly explores doctors' personal perspectives on this topic. This highlights the need for an in-depth investigation into the personal

experiences of doctors in relation to CF, and in particular their views on the impact of culture.

This study seeks to answer the following research questions: "What are the triggers for doctors experiencing Compassion Fatigue?" "What are the coping strategies used by doctors to manage their emotional well-being?", and "How do culture and medical training impact the experiences of Compassion Fatigue amongst doctors?" These questions are explored using semi-structured interviews with doctors. This study aims to improve understanding of the mental health problems doctors face, better inform the development of strategies to reduce CF, and improve patient and staff well-being.

■ Literature Review

Currently, the literature is full of inconsistently defined terms. These terms, including burnout, CF, and vicarious trauma, share similar symptoms, resulting in unclear distinctions between them. CF was chosen as the key term in this research as it uniquely captures the emotional cost of caring for others, which many healthcare professionals navigate daily. However, it should be noted that the terms "burnout" and "vicarious trauma" are often used interchangeably with the term CF in the literature, and therefore, research that focuses on these constructs is included in the current review.

A range of individual and organizational triggers has been identified in the literature as contributing to the development of CF in doctors. Exposure to distressing events, including young or traumatic deaths, cases physicians can relate to, and

witnessing the impact of death on patients' relatives, has been highlighted as impactful.⁵ Howard *et al.* discovered that these events were further intensified by the burden of persuasive responsibility, fear of making mistakes, and conflict in the workplace.⁵

However, the literature consistently highlights that CF is also driven by routine events and heavy workloads.^{1,2} Small, seemingly insignificant incidents accumulate over time, gradually eroding resilience and increasing vulnerability to stress. These pressures are exacerbated in public healthcare systems, where long hours and heavy workloads predict emotional exhaustion.⁹ Research indicates that CF is triggered by the combined effects of acute emotional demands, cumulative routine stressors, and organizational pressures, particularly in contexts characterized by limited institutional support and workplace disempowerment.⁹

Whilst a range of coping strategies and organizational support for doctors' well-being have been identified, substantial barriers to their use remain. Whitebird *et al.* found that healthcare workers experiencing moderate to high stress commonly used coping strategies such as leaning on supportive relationships, engaging in physical activities, and meditation exercises.¹⁰ In particular, working in a supportive team reduces isolation by sharing the burden of care.¹⁰ Healthcare organizations also used certain tactics to promote the well-being of their workers, including encouraging self-care and organizing support groups for stress and bereavement situations.¹⁰

Nevertheless, previous research has shown that there were punitive outcomes for physicians acknowledging and seeking treatment for mental health problems, including stigma, censure by licensing boards, and removal of patient care practice privileges.¹¹ Therefore, despite these strategies being promoted in theory, doctors are discouraged from pursuing them in practice.

There has been a recent shift in the focus of research from triggers and coping strategies to how individual perception of compassion shapes experiences of CF. Most current research is based on the assumption that compassion is a finite resource that can run out.¹² However, Gainsburg and Cunningham discovered evidence to suggest that mindsets about whether compassion is limited or unlimited impact people's experiences and behaviors when compassion is required of them.¹³ Their work showed that not only did participants with limited compassion mindsets self-report negative associations with compassion, but in experimental conditions, they provided lower-quality social support.¹³

Furthermore, Roskvist and Consedine demonstrated that perceptions of compassion are malleable and can be made either more positive or more negative depending on how the concept was framed to participants before they were asked to rate their views of compassion.¹² The participants exposed to a positive framing rated compassion as more satisfying and pleasurable, while the group exposed to a negative framing rated compassion as more draining and exhausting.¹² These pieces of evidence raise the question of whether differences in cultural background and medical training could influence doctors' un-

derstanding of compassion and its impact on their experiences of CF.

Although there is existing literature on the impact of culture/race on the rate of burnout in healthcare professionals, studies have obtained mixed evidence on this matter. Some studies have found that resident physicians in minority racial ethnic groups were more likely to experience burnout,¹⁴ while other studies found the opposite or no difference.^{15,16} It has been argued that the conflicting findings on burnout and CF across racial/ethnic groups may reflect qualitative differences in experience that are not consistently or accurately captured by existing measures. For example, research suggests that minority experiences may be shaped by unique stressors or by differences in how such experiences are perceived and reported compared to majority ethnic groups.^{13,17} This topic is a crucial area to research for clearer answers, so qualitative semi-structured interviews were conducted in this study with physicians from a range of cultural backgrounds. This approach enables more in-depth and nuanced data to be collected and analyzed, in an attempt to bridge this gap in the literature.

■ Methods

This study was conducted through semi-structured interviews, aimed at answering the research questions:

1. What are the triggers for doctors experiencing compassion fatigue?
2. What are the coping strategies used by doctors to manage their emotional well-being?
3. How do cultural background and medical training impact the experiences of compassion fatigue amongst doctors?

Recruitment:

Participants were recruited through opportunistic and snowball sampling. Emails were sent to potential participants who were known to/referred to the researcher through a range of personal, academic, and community-based connections.

Consent:

Before interviews were scheduled, participation information and consent forms were sent to all participants for review and signature.

Data Collection:

The participants took part in a semi-structured interview where they were asked to discuss their experiences of CF. The semi-structured approach allowed a set of predetermined topics to be examined, while also allowing new emerging themes to be explored through follow-up questions to participants' initial responses. This flexible approach allowed for nuanced data to be collected, which cultivates a richer understanding of the participants' experiences. All of the interviews were conducted online via Google Meet, recorded, and manually transcribed by the author. The average duration of each interview was 27 minutes.

Data Analysis:

A hybrid approach was used to thematically analyze data collected from interviews. This approach allowed for the identification of themes grounded in the data, while also drawing on ideas present in the existing literature. The researcher first familiarized herself with the data by manually transcribing the interviews and engaging in multiple read-throughs of the transcripts. An initial coding framework was developed based on the existing literature on CF to guide the initial line-by-line coding of the transcripts. The framework, however, remained flexible to the addition of inductive codes that reflected the patterns seen in the data set. This formed part of an iterative process involving moving back and forth between the initial coding framework and the data. The codes were then organized into a broader set of a priori (deductive) and posteriori (inductive) themes based on patterns of shared meaning.²⁰ Reflexive notes were taken throughout this process to track and review coding decisions. Direct quotes were chosen from the transcripts to illustrate the different themes throughout the results section.

■ Results

Participants:

Seven participants were interviewed. All participants were male and had a mean age of 43.4 years (range 27-59 years). Participants worked across a range of specialties, with professional experience ranging from 3 to 31 years (mean = 18.6 years). Participants came from a range of cultural backgrounds. All had trained or practiced in the UK during their careers, while some had also trained or practiced in other countries (Table 1).

A priori Themes:

The a priori themes, which were identified before data analysis, were derived from the existing literature and research questions. These themes provided the initial framework to guide analysis.

Triggers of Compassion Fatigue:

Exposure to distressing events at work:

Several clinical events were described as emotionally disruptive and deeply impactful for participants. The events included: death of patients, breaking bad news, assault, and suicide. Dr. 1, an intensivist, noted that the death of younger patients impacted him. He described these incidents as very overwhelming: "If it's a patient who's younger, if it was a kid or, like, the younger they are, the more you would feel".

For surgeons and intensivists, breaking bad news to patients and relatives was identified as particularly difficult. "Speaking to the patient, speaking to their family, it's really, really difficult. It takes a lot of emotional energy from you," Dr. 2 explained.

Assault was identified as a significant factor influencing physicians' levels of compassion, with some participants describing it as a normalized aspect of their work. Dr. 5, a psychiatrist, recollected, "I've been assaulted, my colleagues have been assaulted in the past, and that negatively affects compassion". He

further described dreading going to work in the days that followed these incidents.

Participants also described patients and colleagues committing suicide, and how this affected their emotional well-being and reduced their capacity to provide compassionate care. In particular, the suicide committed by a colleague was described as particularly emotionally disruptive. Dr. 5 and 7 noted difficulties providing care while processing the tragedy. Dr. 7, an anesthesiologist, shared,

"We did have one of our colleagues take his own life about 5 years ago during COVID, and he was a chap from the Philippines, and he was away from his family and obviously finding it too much, but there were no real warning signs that anyone was particularly aware of. And it was utterly shocking at the time, and it really had a big impact on a lot of people for a while."

Heavy workload:

Participants frequently described substantial workloads and demanding schedules as factors that diminished their capacity for compassion towards their patients. Many doctors reported working in understaffed hospitals, where they are expected to shoulder the workload of several physicians. Some participants expressed regret at being unable to provide more emotional support for their patients. Dr. 1, an intensivist who worked in the NHS, said, "you're sort of doing the bare minimum for patients because you don't have enough time in the day to focus too much on each of them ... there's just so much to do as a doctor, there's an unending amount of work". Similarly, Dr. 4, a urology surgeon working in the NHS, said he is "basically overwhelmed with the amount of work and the amount of responsibilities".

Thinking about patients non-stop:

Participants also described thinking about patients too much as a common trigger of compassion fatigue. Dr. 3, an orthopedic surgeon, said, "If you do any complex case, you are always thinking about that complex case, till everything completely becomes normal [and the patient has recovered]". Dr. 5, a psychiatrist, explained that "you'll get used to it, but you'll get used to it in a bad way, and that's not good for you, and it's not good for your mental health, and it's not good for your patients. There needs to be times when you need to step away from them".

Coping Strategies:

Participants identified several coping strategies that they used to cope with the emotional toll of practicing medicine. A range of methods was used depending on the nature of the events witnessed or experienced.

Building resilience:

All participants described an expectation within medicine that doctors should build resilience to the emotional toll of clinical practice over time. Although they developed other coping strategies, resilience was noted as a fundamental part of clinical development. Dr. 6, a psychiatrist, reported that it was

seen as “just part of the training, you build resilience and over time, [you] just have tough skin”.

Emotional barriers:

Many doctors noted that they create emotional barriers to protect themselves from the emotional toll of their work. Dr. 4, a urologist, explained that “It's not a matter of being insensitive or uncompassionate, but it's just like a coping mechanism to basically carry on with your daily life”. Dr. 2, another urologist, stated, “It can be overwhelming otherwise” if there are no barriers in place.

Talking to other medics:

Participants with friends or family members working in medicine emphasized the importance of being able to talk about work-related stresses with those who understand the profession. Similarly, those without friends and family in the profession identified supportive colleagues as crucial to helping them cope. “I've got really nice colleagues, and we talk all the time ... That's quite important”, Dr. 2 said. Several other participants agreed, expressing how this sense of camaraderie among colleagues has facilitated their learning and supported them emotionally. Dr. 4 also noted that “usually, medical people understand each other more than other people”, so he talks a lot with his friends and his wife, who all work in medicine.

Signs of Compassion Fatigue:

Disturbed sleep:

Several participants described not being able to sleep well after working long shifts. Dr. 3, an orthopedic surgeon, said, “When you are sleeping, you suddenly have a dream ... that patient or that procedure will come to you”. This seems to be the case not only for surgeons but also for psychiatrists. Dr. 6 said, “You can't sleep with your two eyes closed until you retire, you can't”. Others also mentioned difficulty with sleep during their rotation in the emergency room in their foundation years due to the severity of the cases witnessed.

Passive desensitization:

Some participants noted that consistent exposure to trauma, regardless of years in the profession, could lead to desensitization and make it increasingly difficult to maintain compassion towards patients. Dr. 3 explained, “The more you're exposed to stuff in medicine, the more effort it takes for you to be compassionate because you get used to the things that you see and experience”. Although desensitization may appear to function as a coping strategy, participants noted that they are becoming desensitized passively over time, which differs from using desensitization as an active strategy to promote coping. Dr. 2 noted that while the concern you have for your patients can be overwhelming at times, “the moment that you're not worried about what you're doing to people and the effect that you have on them is probably a time to try and get some help for yourself”.

Influence of Cultural Background on Compassion Fatigue:

Cultural variations in the stigmatization of mental health:

Participants born and raised outside of the UK shared the view that there was less mental health awareness in their home countries. They described a lack of support for doctors following traumatic events. Dr. 4 recalled, “Back in my country, really catastrophic things happen, and no one really takes care of that aspect of things; it really takes a toll and keeps a scar in people's minds”. Several participants reported that speaking out about personal struggles is frowned upon in their home countries, where mental health is still very much stigmatized. Dr. 6 noted that “there is still some stigma [in Nigeria] ... because people are not speaking up.” While Dr. 4, who is from Lebanon, stated, “I come from an eastern country where our feelings are looked upon as weakness rather than reality”. Dr. 1, who is practicing in Hong Kong, also commented that “the culture, particularly in Asia overall, not just Hong Kong, there's sort of a stigma behind mental health, and it's not talked about too much. It's talked about a lot more these days, but still not as much as in Western counterparts”. Contrastingly, doctors from the UK reflected on what they perceived as a growth in the prioritization of mental well-being.

Reduced support systems for immigrants:

Immigrating to other countries impacted the support available to participants. Dr. 6, who immigrated from Nigeria to the UK, noted the lack of support he has now compared to when he lived in Nigeria: “Back home you've got family support, there's community support. I don't really have anyone ... you may end up realizing the person you want to speak to is miles away from you”. He further explained that seeking support in an unfamiliar country can be particularly difficult, as people who are not connected to your culture or background may find it more difficult to relate to your struggles and experiences.

Compassion as a Limited and Fluctuating Resource:

All participants spoke about compassion as a limited and fluctuating resource. For example, Dr. 2 noted its finite capacity, stating that “we've all got a certain amount of bandwidth that we can expend on things at work, on ourselves, on our family ... probably there is some compassion exhaustion for some people when they've got a lot of other stuff going on”. While Dr. 7 agreed that compassion can run out, he noted that this occurs only during “brief moments, particularly during shifts, and particularly when you're a resident doctor where you're doing a row of night shifts ... because you've just completely run out of energy”. Dr. 5 suggested that fluctuations in compassion is something we have control over, believing that “you can influence [your level of compassion] without a doubt, and you can always make it better than it was yesterday or worse than it was yesterday”. How one can influence compassion in this way, however, was not established during the interview.

Posteriori Themes:

The posteriori themes were identified inductively from the data during analysis.

Lack of Preparation in Medical School:

All participants reported receiving little to no education during medical school on coping with the stresses of the profession. As a result, they had to develop their own coping strategies. Dr. 3, who trained in India, said that “even now, I’m not sure whether the medical degree in the UK or even in India they teach you how to cope with the failures”. This experience was consistent across participants, regardless of differences in attitudes towards mental health in the countries in which they were trained. All participants indicated that having some support during medical school would have helped them a lot.

Attitudes About Mental Health:

Differences across specialties:

Several participants across different specialties noticed differences in how their colleagues viewed mental health. In particular, it was not seen as unusual in psychiatry and anaesthetics to check on each other’s mental well-being. Dr. 5, a psychiatrist, said “it’s not expected, but it’s not unusual to always check in with people and to always talk about mental health”, while Dr. 7, an anesthesiologist, said “you can always find someone who you’re able to talk to and there’s often an office, you can go in and have a chat with [colleagues] about what’s happening, and most people can do that fairly openly”. He thinks this is the case as “the suicide rate amongst anesthesiologists is one of the highest”, so they are willing to be more open in an attempt to reduce the rate of self-harm. On the other hand, several participants noted that there is much more stigma in other specialties, particularly General Practice. However, the reason for this difference was not established in the interview.

Improving attitudes about mental health in the UK:

On a more hopeful note, the participants currently working in the UK observed a positive shift in people’s views on mental health. Dr. 7 noticed that “there’s a much lower threshold to take time to manage your mental health, to seek help, and to alert people when you are feeling that work is getting to you and that you are struggling”. Dr. 4 also noted how “people [in the UK] are more in touch with what they feel, and they’re a bit more open to share.”

Triggers of Compassion Satisfaction:

Multiple participants identified positive patient feedback as a primary motivator for continuing to practice medicine. Dr. 3 explained, “When you see the patient coming back to your clinic saying that, well, you changed my lifestyle ... that makes you happy, and that is like fuel for your motivation”. Dr. 6 described a similar experience, stating, “When I see my patients get better, when I see them recover, when I see them happier than when they came to me, and most importantly, when I see them being productive as well. That is fulfilling”. Similarly, Dr. 1 felt motivated as he knew that “you’re doing a lot of good, you’re changing a lot of lives”. He currently works in the ICU and noted that seeing “how the family of the patients shows so much happiness and gratitude after you’ve saved their lives” gave him a lot of satisfaction.

■ Discussion

Discussion

In relation to the first research question, the findings suggest that doctors’ experiences of CF were triggered by a range of clinical events and organizational pressures. Exposure to distressing events was also identified as a key trigger in research conducted by Howard *et al.* and Dasan *et al.*, respectively.^{3,5} The findings of the current study build on these findings and additionally suggest that the types of distressing events perceived to trigger CF vary somewhat across medical specialties. This could be linked to the differing nature of the types of cases and the role clinicians play in a patient’s care within each specialty. For example, a large part of surgeons’ jobs includes talking with patients about their diagnoses and treatment plans; therefore, breaking bad news was perceived as a major trigger of CF and emotional exhaustion for these participants. These findings highlight how the emotional demands associated with CF may differ according to specialty-specific clinical experiences.

Limited research has been conducted on the impacts of CF on physicians in different specialties. The current findings raise important questions about the different emotional needs of different specialists, as well as how they could be better supported with suitable care and prevention. In addition, an emergent theme in the data was the perception that attitudes towards mental health differed across specialties, with some having higher levels of awareness and support than others. This implies that interventions aimed at reducing CF may benefit from being tailored to the culture and demands surrounding each particular specialty.

The current study also found that certain organizational factors, such as high-intensity workloads, could trigger CF in participants by limiting their emotional capacity to engage compassionately with patients. This was particularly evident for participants when they were at the start of their careers, when their workload was at its heaviest. These findings support previous research by Dasan *et al.*, who showed that doctors’ levels of compassion are reduced when they are frequently exhausted during their training years.³ Their study also reported that younger consultants reported higher levels of CF, whereas older consultants reported lower levels of CF.³ A similar pattern emerged in the current study. Consultants more often described experiencing compassion satisfaction triggered by receiving positive patient feedback, while younger doctors focused more on the physical and emotional exhaustion of their work.

While these findings identify workload as a key trigger of CF, they also address the second research question by suggesting that doctors may manage CF through the development of resilience over the course of their professional experience. These findings may indicate that early-career doctors may be particularly vulnerable to CF due to the combination of intensive workloads and limited experience with handling the intense emotional demands of the role. In contrast, greater emphasis on compassion satisfaction amongst consultants may suggest that some doctors become more able to manage these demands over time, through more manageable workloads and

resilience-building through experience. However, it is important to note that in both the current study and Dasan *et al.*'s study, survival bias may have played a major part in forming this pattern.³ Those who struggled to cope during residency may be more likely to leave medicine and pursue another career, while those who coped well may be more likely to pursue consultancy roles. To eliminate this possibility, further research utilizing a longitudinal design to follow the trajectories of doctors' careers and their experiences of CF is required.

The lack of training participants received to prepare them for the emotional burden associated with clinical practice may have contributed to the CF they described experiencing within their roles. Despite participants in the current study suggesting that mental health awareness has improved among doctors in the UK, it is evident that improvements are still needed, as this appears not to have translated into tangible support during training. Given that the findings indicate younger doctors struggle more with coping with the stress of the profession, mental health training should be incorporated into the curriculum in medical school. Future research should consider the current barriers to providing relevant training in medical school, given that it has been noted as an issue by participants who had trained in different countries. This could better inform the best form of support to provide.

It is also notable that doctors understood that compassion fluctuated over time, with some even suggesting that these fluctuations could be controlled by the individual. Previous research by Roskvist, Consedine, Gainsburg, and Cunningham supports this finding, demonstrating how perceptions of compassion are malleable and can positively or negatively impact people's experiences.^{12,13} This suggests that CF is not inevitable and could be positively influenced by appropriate interventions, further emphasizing the importance of training and support for doctors to help them manage their mental well-being.

In relation to the third research question, the findings indicate that participants' experiences were shaped by the varying levels of mental health awareness and stigma in which they trained and worked. Some participants have trained in a culture of higher levels of stigma around mental health compared to the UK. This may have resulted in reluctance to seek help for mental health difficulties, which aligns with Garcia *et al.*'s conclusion that stigma may reduce disclosure and help-seeking in minority racial/ethnic groups experiencing mental health concerns.¹⁹ In addition, these doctors may have less access to certain coping strategies, such as social support. Dr. 6's comments about not having anyone to talk to and Lawrence *et al.*'s findings support this interpretation, suggesting that doctors coming to the UK from other cultures may be more vulnerable to CF.¹⁶ This lends some support to previous research concluding that resident physicians in minority racial ethnic groups were more likely to experience burnout.¹⁴ As a result, confidential support from external professionals may be helpful for them.

Despite these findings, participants offered limited additional insights into how cultural values shaped people's understanding of compassion and CF. This could suggest that participants may not always be consciously aware of cultural factors that

may shape their experiences of CF. This is significant as direct questions about culture did not elicit many insights, implying that participants may not be consciously applying cultural values in their compassionate practice. Nevertheless, this does not rule out unconscious influence. Furthermore, discussion relating to mental health stigma and how changes in support systems impacted CF was mainly driven by participants born outside of the UK. This suggests that cultural influences became more noticeable when participants were able to compare experiences across different cultural systems.

■ Conclusion

This study explores experiences of CF in seven doctors representing a wide range of clinical specialties, cultural backgrounds, and years of experience in the field. The findings of this study built on and helped to corroborate the findings of existing studies, which have investigated the experiences of CF in doctors. Similar triggers and coping strategies were identified in the current study to those in the existing literature. This study has further extended understanding by examining the impacts of culture and medical training on CF. Preliminary evidence was found that suggested both the cultural backgrounds doctors came from and the culture within their specialties influenced their attitudes towards mental health, thus affecting their willingness to seek help when needed. The study concluded that it may be helpful to offer more support to physicians, particularly those who emigrated to the UK, as they may lack social networks and may come from cultures where mental health is heavily stigmatized. Finally, this study identified differences in CF experiences by participants' specialisms and length of time spent in the profession. These findings suggest that younger doctors should be better supported in the early years of training and practice, and interventions could be tailored more specifically to fit the cultures surrounding different specialisms in order to be the most effective. In both cases, implementing structured peer support programs may prove useful in supporting doctors in their work. Without these interventions, the rate of burnout and CF may continue to increase, reducing doctor retention rates.

■ Limitations

While the current study offers some helpful insights into the field of CF, there are several limitations to the study that ought to be acknowledged when interpreting the findings. Difficulties in obtaining access to suitable participants in this study led to a small, all-male sample, significantly limiting the generalizability of the findings. While the study offers interesting preliminary findings, further research with larger, mixed-gender samples is needed to substantiate them. In particular, it is important to recognize that the absence of female participants means the results do not contain insights from female doctors, who may have a range of different experiences with CF.

In addition, while interviews provide valuable insight into personal perspectives, they are insufficient for capturing culturally embedded beliefs that are not easily articulated through self-report. Participants may not have been consciously aware

of such beliefs, limiting the insights possible using this methodology. Methods allowing deeper exploration of less conscious thoughts and opinions are needed for further research.

Furthermore, it is relevant to highlight that all the interviews were conducted by a high school student, and the significant difference in age and experience between the researcher and participants may have resulted in participants being hesitant to go into detail around sensitive topics. This may have been born of a desire not to pass on emotional burden, or of discomfort with revealing details to someone not in the same field. While it is important to acknowledge this as a potential limitation, there were attempts to minimize the power differential through building rapport and collecting informed consent from participants before interviews were conducted.

■ Further Research Directions

The current study demonstrates that experiences of CF vary across specialties and cultural backgrounds. These results reveal areas where further research is needed, as the data set for this study is very limited. For example, questions raised include: how can doctors in training be best supported in coping with the emotional toll of practicing medicine, and what kinds of support can be provided to doctors practicing outside of their home countries? Longitudinal studies could be conducted to inform the development of support programs for training doctors. Furthermore, focus group discussions, where participants are able to discuss their experiences and opinions through guided conversation, may be a more fruitful method of data collection on how culture influences one's thought patterns and actions.²¹ This method allows participants to build on each other's ideas, which may be more effective in capturing implicit data such as shared assumptions and unspoken norms.

■ Acknowledgements

The authors would like to express their deepest gratitude to the participants for their time, insights, and willingness to share their experiences; without them, this study could not have been completed. Special thanks are due to Mrs. Hazel Knight, Head of Research at the Perse School, for facilitating introductions to key participants, which were vital for the data collection phase of this study.

■ References

1. Medical and Dental Defense Union in Scotland. *Compassion fatigue in healthcare professionals*: Glasgow, Scotland, 2025. Compassion-fatigue-report.pdf (accessed 22/12/2025).
2. Figley, C.; Huggard, P.; Rees, C. E. *First do no self-harm: understanding and promoting physician stress resilience*; Oxford University Press: New York City, 2013.
3. Dasan, S.; Gohil, P.; Cornelius, V.; Taylor, C. Prevalence, causes and consequences of r and compassion fatigue in emergency care: a mixed-methods study of UK NHS Consultants. *Emergency Medicine Journal* **2015**, *32* (8), 588–594. <https://doi.org/10.1136/emmermed-2014-203671>.
4. Wells-English, D. Compassion fatigue and satisfaction: Influence on turnover among oncology nurses at an urban cancer center. *Number 5/October 2019* **2019**, *23*(5), 487–493.
5. Howard, L.; Wibberley, C.; Crowe, L.; Body, R. How events in emergency medicine impact doctors' psychological well-being. *Emergency Medicine Journal* **2018**, *35* (10), 595–599. <https://doi.org/10.1136/emmermed-2017-207218>.
6. Kariyawasam, L.; Ononaiye, M.; Irons, C.; Kirby, S. E. A cross-cultural exploration of compassion, and facilitators and inhibitors of compassion in UK and Sri Lankan people. *Cambridge Prisms Global Mental Health* **2022**, *9*, 99–110. <https://doi.org/10.1017/gmh.2022.10>.
7. Lawn, S.; Roberts, L.; Willis, E.; Couzner, L.; Mohammadi, L.; Goble, E. The effects of emergency medical service work on the psychological, physical, and social well-being of ambulance personnel: a systematic review of qualitative research. *BMC Psychiatry* **2020**, *20*, 1–16.
8. Crowe, L. Identifying the risk of compassion fatigue, improving compassion satisfaction, and building resilience in emergency medicine. *Emergency Medicine Australasia* **2016**, *28* (1), 106–108. <https://doi.org/10.1111/1742-6723.12535>.
9. Welp, A.; Meier, L. L.; Manser, T. Emotional exhaustion and workload predict clinician-rated and objective patient safety. *Frontiers in Psychology* **2015**, *5*, 1573. <https://doi.org/10.3389/fpsyg.2014.01573>.
10. Whitebird, R. R.; Kreitzer, M.; Crain, A. L.; Lewis, B. A.; Hanson, L. R.; Enstad, C. J. Mindfulness-based stress reduction for family caregivers: a randomized controlled trial. *The Gerontologist* **2013**, *53*(4), 676–686. <https://doi.org/10.1093/geront/gns126>.
11. Saddawi-Konefka, D.; Moutier, C. Y.; Ehrenfeld, J. M. Reducing Barriers to Mental Health Care for Physicians: An Overview and Strategic Recommendations. *JAMA* **2025**.
12. Roskvist, R.; Consedine, N. S. Is compassion fatigue a self-fulfilling prophecy in health care? A preregistered experimental study of manipulated expectations in doctors and medical trainees. *British Journal of Health Psychology* **2025**, *30* (3), e70013. <https://doi.org/10.1111/bjhp.70013>.
13. Gainsburg, I.; Cunningham, J. L. Compassion Fatigue as a Self-Fulfilling Prophecy: Believing compassion is limited increases fatigue and decreases compassion. *Psychological Science* **2023**, *34* (11), 1206–1219. <https://doi.org/10.1177/09567976231194537>.
14. Fang, D.; Young, C.; Golshan, S.; Moutier, C.; Zisook, S. Burnout in premedical undergraduate students. *Academic Psychiatry* **2012**, *36* (1), 11–16. <https://doi.org/10.1176/appi.ap.10080125>.
15. Meredith, L. S.; Hackbarth, N. S.; Darling, J.; Rodriguez, H. P.; Stockdale, S. E.; Cordasco, K. M.; Yano, E. M.; Rubenstein, L. V. Emotional exhaustion in primary care during early implementation of the VA's medical home transformation: Patient-aligned Care Team (PACT). *Medical Care* **2015**, *53*(3), 253–260.
16. Lawrence, J. A.; Davis, B. A.; Corbette, T.; Hill, E. V.; Williams, D. R.; Reede, J. Y. Racial/ethnic differences in burnout: a systematic review. *Journal of racial and ethnic health disparities* **2022**, 1–13.
17. Maslach, C. *Burnout: the cost of caring*; Ishk: Los Altos, California, 2003.
18. Dyrbye, L. N.; Burke, S. E.; Hardeman, R. R.; Herrin, J.; Wittlin, N. M.; Yeazel, M.; Dovidio, J. F.; Cunningham, B.; White, R. O.; Phelan, S. M.; Satele, D. V.; Shanafelt, T. D.; Van Ryn, M. Association of clinical specialty with symptoms of burnout and career choice regret among US resident physicians. *JAMA* **2018**, *320* (11), 1114. <https://doi.org/10.1001/jama.2018.12615>.
19. Garcia, L. C.; Shanafelt, T. D.; West, C. P.; Sinsky, C. A.; Trockel, M. T.; Nedelec, L.; Maldonado, Y. A.; Tutty, M.; Dyrbye, L. N.; Fassiott, M. Burnout, Depression, Career Satisfaction, and Work-Life Integration by Physician Race/Ethnicity. *JAMA Network Open* **2020**, *3* (8), e2012762. <https://doi.org/10.1001/jamanetworkopen.2020.12762>.

20. Swain, J. *A hybrid approach to thematic analysis in qualitative research: Using a practical example*, 2018. <https://doi.org/10.4135/9781526435477>.
21. Morgan, D. L. (1996). Focus groups. *Annual review of sociology*, 22(1), 129-152.

■ Author

Lea is an A-Level student at the Perse School, Cambridge, with a passion for medicine. She hopes to read Medicine at a UK university, as well as contribute to the field of medical research in the future.

■ Appendix

Informed Consent Form

Please read the following information sheet, which details information about this research project and what participation in this project will involve. Afterward, if you would like to participate, you will be asked to sign and return the consent form. This consent form is necessary for us to ensure you fully understand the purpose of your involvement and that you agree to the conditions of your participation.

Research project title: Experiences of compassion fatigue and compassion satisfaction in hospital doctors: an exploration of triggers, coping strategies, and the impacts of doctors' cultural backgrounds.

Student researcher: Lea Yuen

Purpose of the project: Due to the prolonged exposure to suffering and trauma healthcare workers face every day, their mental health is inevitably negatively impacted, with a high percentage having reported experiencing compassion fatigue as a result of their job. Figely (2002) defined compassion fatigue as a reduction in the capacity or interest in bearing the suffering of others. It has also been recognized, however, that individuals who work in such conditions can also derive pleasure from doing their work well, defined by Stamm (2010) as compassion satisfaction. Research (Dasan, 2015) indicates that compassion fatigue in medical professionals can lead to stress, burnout, and a loss of compassion, resulting in poor quality care. Thus, this research aims to better understand how hospital doctors experience compassion fatigue and compassion satisfaction, with the aim of helping to inform the development of strategies to reduce compassion fatigue and promote compassion satisfaction. In particular, this research seeks to explore how cultural backgrounds may influence doctors' experience of compassion fatigue, as currently there is a lack of research examining in-depth self-reports of their experiences in this area.

What will I be asked to do?

I am asking for your voluntary participation in an interview as part of my academic research project. The interview will take around 30-40 minutes, but may be much shorter, and will take place on Google Meet at a time most convenient for you.

Risks and Benefits Associated with Participation:

There are no anticipated risks associated with your participation, but it is possible that you may feel some emotional discomfort about the questions asked. If this occurs, you have the right to stop the interview, withdraw from the research at any point, or not answer any specific question. You will be given

an outline of the interview questions prior to the interview, which will help you decide if you would like to participate or not.

You will receive no direct benefit from participating in this research, but this research will hopefully contribute to existing literature to deepen our understanding of this topic and improve mental health for doctors.

What will happen to my data?

- The interview will be recorded, and an anonymized transcript will be produced for data analysis.
- You will be sent the transcript and given the opportunity to correct any factual errors if you wish.
- The transcript will be analyzed by student researcher Lea Yuen.
- Access to the anonymized transcript will be limited to Lea Yuen and academic colleagues and researchers with whom she might collaborate as part of the research process.
- Any summary interview content, or direct quotations from the interview that are included in an academic publication, will be anonymized so that you cannot be identified, and care will be taken to ensure that other identifying information is not revealed.
- The actual recording will be destroyed after the transcript is produced and approved by all participants.
- Any variation of the conditions above will only occur with your further explicit approval.

By signing this form, I agree that:

- I am voluntarily taking part in this project. I understand that I do not have to take part, and I can stop the interview at any time;
- I may be directly quoted from the transcribed interview with a pseudonym (made-up name) in the published research;
- I do not expect to receive any direct benefit or payment for my participation;
- I can request a copy of the transcript of my interview and may make edits I feel necessary to ensure the effectiveness of any agreement made about confidentiality;
- I have been able to ask any questions I might have, and I understand that I can contact the researcher with any questions I may have in the future;
- I have read all of the information above.

If you have any questions about this study, please feel free to contact:

Student Researcher: Lea Yuen

Email: lyuen@perse.co.uk

or

Research Mentor: Mrs. Hazel Knight

Email: hkknights@Perse.co.uk

Printed Name _____

Participant's Signature _____

Date (dd/mm/yy) _____

Media Sentiment and Economic Signals of Emerging Blockchain Adoption in Vietnam

Gia Minh (Michael) Pham

The Hill School, 860 E High Street, Pottstown, PA 19464, USA; minhphamgia889@gmail.com

ABSTRACT: This paper examines the emerging relationship between blockchain-based media sentiment and relevant economic indicators in Vietnam. Although Vietnam is positioning itself to become a regional leader in blockchain technology, frameworks for adoption remain under development and are ambiguous. Also, due to a lack of enterprise-driven data, no studies to date have formally assessed its impact on the economic landscape. Addressing this gap, we construct a multi-layered approach, integrating quantitative (news sentiment and six macroeconomic indicators) and qualitative data (policy contexts and two technology readiness indexes). Using natural language processing (NLP) models, we analyze sentiment across thousands of Vietnamese-language articles from four popular newspapers, filtered by blockchain-related keywords. We then align sentiment trends with macroeconomic variables in a time-series perspective to perform EDA, exploring whether fluctuations in media sentiment correlate with economic shifts. While causality is not definitive, findings show meaningful correlations between indicators that are congruent with qualitative factors. Conclusively, our research methodology aims to lay the foundation for evaluating economic signals of developing countries at an early stage.

KEYWORDS: Development Economics, Blockchain Adoption (BCA), Macroeconomic Indicators, Sentiment Analysis, Natural Language Processing (NLP), Regression, Granger Causality.

■ Introduction

The emergence of blockchain (BC) technology is increasingly prominent, serving as a remedy for prolonged infrastructure deficits and a catalyst for the efficacy of diverse industries in emerging economies. However, insufficient public data and barriers to adoption present a challenge to drawing empirical conclusions at an early stage. Our multidimensional analysis, both dynamic and exploratory, investigates how BC diffusion relates to economic development, aiming to foster trust and efficiency in the Vietnam BCA framework. In this paper, we answer three core questions:

- RQ1: How is sentiment measured quantitatively in our dataset?
- RQ2: How do shifts in Vietnamese media sentiment associate with BC-related economic activity?
- RQ3: Which macroeconomic indicators exhibit significant relationships with blockchain adoption sentiment in Vietnam?

The remainder of this paper proceeds as follows. The background section provides the context of the paper, portraying the BC situation in the ASEAN and Vietnam contexts. The literature review section reviews existing papers on related works. We then present our methodology section, detailing data collection, sentiment-scoring models, and time-series techniques. In the findings section, we use correlation, regression, and Granger causality to identify relationships between news sentiment and economic signals. In the following section, we discuss the study's limitations and directions for future research. The paper concludes with a summary of important insights.

■ Background

Amidst rapid changes in technological innovation, the influence of emerging markets in the global economy, especially those situated in Southeast Asia, has overshadowed the growth of advanced economies. Characterized by burgeoning financial infrastructure and evolving institutional frameworks, they garner much attention from practitioners and investors alike. Their growth trajectory fuels capital flows, trade, and economic prosperity. However, it also presents unique challenges. In many Association of Southeast Asian Nations (ASEAN) countries, financial inclusion is notably uneven. For example, approximately 98 million Indonesian people (48% of the country's adult population) remain underbanked, restricting access to savings, credit, and basic payment services.¹ As a high proportion of SMEs constitutes the private sector, opaque supply chains are increasingly susceptible to market trust and quality assurance issues. Most logistics of goods and services are operated without verification.²⁻⁴ In addition, record-keeping in public services is largely paper-based and bureaucratic. Such is the underlying circumstance in Thailand, where centralized databases suffer from inefficiency and a lack of transparency.⁵ Banking systems and business corporations experience similar centralization, resulting in relatively low public confidence across institutions.⁵ Collectively, these gaps in financial inclusion, supply chain vulnerability, governance structure, and transparency hamper economic resilience and pose barriers to artificial intelligence (AI) and digital transformation.

A promising solution to emerging markets' challenges is blockchain. The term, which originated from Satoshi Nakamoto's 2008 Bitcoin whitepaper, one of the earliest forms of BC

applications, is a distributed ledger technology that addresses structural challenges securely and transparently.⁶ Fundamentally, BC records a sequence of blocks using cryptography, where each block—containing transaction records—is linked to the previous block. One of its key attributes is immutability. Once a block is added to the ledger, the block order is irreversible and cannot be modified, creating a tamper-proof audit trail for corporate registries and end-to-end provenance of goods.^{6,7} To bolster data integrity, the BC system is built on a peer-to-peer (P2P) network of consensus mechanisms, allowing individuals to perform trustless transactions without the need for central intermediaries, so-called decentralization.^{6,8} On top of that, BC is highly interoperable. It serves as a backbone for seamless exchange of standardized data and automating complex tasks such as credit disbursement, significantly reducing manual processing and intervention, compliance costs, and delays.^{6,7} Thus, blockchain's versatile architecture offers a wide range of practical applications.

By far, digital wallet platforms and decentralized finance (DeFi) represent leading applications of BC technology. Whereas digital wallets incentivize the unbanked population to gain financial flexibility, bypassing conventional banking constraints and transaction fees, DeFi leverages smart contracts to radically transform the financial service landscape, providing automated P2P lending and novel asset tokenization.^{8,9} Yet, the potential of BC hinges on a proper regulatory environment.^{10,11} Conforming to safeguards, the Monetary Authority of Singapore has implemented a stringent sandbox to experiment with integrating BC in Fintech.¹² The Securities and Exchange Commission (SEC) of the Philippines, by contrast, has issued regulations that acknowledge the increased digital asset space.¹³ Despite proactive actions, many governments struggle to keep pace as digital advancements outstrip policy frameworks. The tension between the promise of BC technology and regulatory limitations in emerging economies creates a pressing need for further research.

Vietnam exemplifies this dynamic as an emerging, low-to-middle-income country with unregulated economic and institutional infrastructures due to scarce digital record-keeping.¹⁴ Because of this, Vietnam has yet to develop a viable banking system, transparent supply chains, and effective public government, reflecting underlying structural challenges of a typical emerging market.¹⁵⁻¹⁷ Concurrently, Vietnam's digital asset adoption rate has seen exceptional growth, ranking among the top five countries within three years consecutively.¹⁸ Over 17 million Vietnamese—the majority in early adulthood—have reportedly been involved in the cryptocurrency market, underscoring a strong participation rate in decentralized applications (dApps), including DeFi and P2P activities. Notably, Vietnam secured over 105 billion USD from blockchain investments from 2023 to 2024.¹⁹ While upward trends reflect public optimism, the reverse is true for policymakers. The decentralized and borderless nature of digital assets in the absence of state regulatory oversight has exposed users to unprecedented threats: cyberfraud, money laundering, tax evasion, and personal data breaches.¹⁹ Furthermore, the lack of regulatory clarity forces BC start-ups, such as the on-chain

liquidity protocol Kyber Network and the play-to-earn (P2E) gaming platform Axie Infinity, to register overseas. Consequently, this results in substantial capital outflows and losses in tax revenue, placing a strain on the national budget. As a result, the government is actively pursuing frameworks to address contemporary infrastructure challenges and strengthen user protections, positioning Vietnam to become a regional leader in BC by 2030.²⁰

■ Literature Review

In recent years, numerous studies have been conducted to bring to light the potential of BCA across domains such as education, health, logistics, financial technology, and supply chain management. We determine three pertinent strands of existing BCA studies that build the foundation of our research: (1) frameworks and determinants; (2) news sentiment; and (3) Vietnam's BC economy. We articulate notable gaps across these strands and position our paper to bridge them.

Frameworks and Determinants:

Frameworks, such as the Technology Acceptance Model (TAM), are used as theoretical models to study BCA. At the same time, determinants are specific factors—e.g., perceived usefulness and perceived ease of use—constructed within these models.

Literature surveys elucidate a comprehensive perspective on BC technology, ranging from its historical background and principles to its areas of application.^{6,7,11,21-23} On the other hand, empirical studies develop granular approaches to identify determinants affecting BCA. For instance, Lou & Li propose a conceptual framework, integrating TAM with Innovation Diffusion Theory (IDT) to survey perceptions of BCA from business managers' perspectives in the Taiwan fintech industry.²⁴ However, the work mainly highlights blueprints for future research due to data scarcity and sampling difficulty. Building on TAM methodology, Nuryyev *et al.* examine how individuals in Taiwanese SMEs in tourism and hospitality perceive the usefulness of cryptocurrency payment, a popular application of BC technology.²⁵ Extending contextualized determinants, they find that behavioral intention is engendered by social influence, strategic orientation, and individual SME characteristics. Malik *et al.* conduct a cross-sectional analysis on Australian organizational behaviors through the lens of the Technology Organization Environment (TOE) model, advising future research to enhance generalization.²⁶ Chittipaka *et al.* refine TOE with a larger sample size across many sectors, shedding light on prospects of the Indian supply chain, improving generalizability, and tackling the diminutive data sampling of previous studies.²⁷ Determinants stimulating the adoption are “relative advantage, trust, compatibility, security, firm's IT resources, higher authority support, firm size, monetary resources, rivalry pressure, business partner pressure, and regulatory pressure”. Beyond organizational and behavioral determinants, Guo *et al.* emphasize the interplay between economic factors—such as inflation and exchange rates—and cryptocurrency adoption, employing various econometric models with panel data.²⁸ Similar studies suggest that mac-

ro-level drivers become more pronounced, supporting the justification of indicators in the methodology section.²⁹⁻³²

News Sentiment:

News sentiment analysis (SA) is an emerging technique that extracts the semantic meaning from text. Its application holds immense benefits for technology adoption and business domains, allowing entities to gauge public perception, predict market behavior, and derive quantitative insights. Early on, Godbole *et al.* developed a two-step system to determine the sentiment weighted by each word and aggregate these values to determine the overall sentiment of a text.³³ Shapiro *et al.* created a more advanced, specialized system for capturing economic news. They experimented with the reliability of sentiment in reflecting market responses to macroeconomic variables such as consumption, output, interest rates, and inflation.³⁴ Khedr *et al.* amalgamate news sentiment with historical prices, predicting stock market behavior with up to 90% accuracy via a naive Bayes algorithm.³⁵ SA is also prevalent in cryptocurrency and BC applications. Coulter peruses SA on thousands of international cryptocurrency articles to underscore certain topics that drive the volatility of the Bitcoin price.³⁶ Caviggioli *et al.* explore how SA on specific tweets can affect the corporate reputations of retail companies that adopt Bitcoin payment.³⁷ Recently, the advent of state-of-the-art NLP models—including Large Language Models (LLMs) such as GPT-4, BERT, and FinBert—has outperformed conventional sentiment paradigms. Their efficacy in capturing nuances facilitates investors in navigating the unpredictable nature of cryptocurrencies and optimizing strategies.³⁸ SA techniques have also been adapted for non-English contexts. Vu *et al.* implement the Vietnamese-language model PhoBERT to classify the polarity of nearly 40,000 financial-oriented articles, providing investors with greater data accessibility and certainty when engaging with markets.³⁹ Although PhoBERT achieves 81% accuracy, there is a weak correlation between news sentiment and stock prices, but a strong link between negative sentiment shocks (i.e., sudden decrease of sentiment) and increased market volatility. A possible explanation is that the Vietnamese stock market is relatively young.

Vietnam's Blockchain Economy:

As an emerging market, Vietnam has attracted growing public and academic interest in BCA due to its potential to support economic development and digital transformation. While papers synthesizing the context of blockchain-based economy are abundant,⁴⁰⁻⁴² empirical studies on domain applications are inadequate.

Existing case studies and surveys are concentrated within the supply chain spectrum—including logistics, agriculture, and food supply.^{43,44} Nguyen *et al.* fill the gap by conducting a study on the potential benefits of BC technology in tackling the limitations of traditional real estate.⁴⁵ Both Pham *et al.* and Nguyen *et al.* explore determinants driving blockchain in journalism and logistics based on Unified Theory of Acceptance and Use of Technology (UTAUT), respectively, as an effort to address the nascent stage of BCA in Vietnam.^{46,47} While

the former omit social influence as a determinant of BCA, the latter identify it as the most salient factor via neural-network analysis. Nevertheless, existing fragmented frameworks and ill-defined guardrails have prevented Vietnam from capitalizing on the pre-emptive phase of Asian technological advancement. The legal stance on cryptocurrency is not receptive to recognizing it as legal tender; instead, it is treated as proprietary, citing concerns that cryptocurrency poses immeasurable risks to consumers and has yet to fulfill the monetary role of the Vietnamese dong.⁴¹ However, this approach may stifle innovation in fintech and BC entrepreneurship, causing Vietnam to lag behind its Asian neighbors by overlooking key benefits of BC, such as removing intermediaries, reducing transaction costs, and generating a potential source of tax revenue.^{40, 42-48} In extreme cases, current government practices may inadvertently incentivize unrecognized digital transfer activities, posing even greater risks to consumers. In addition to the absence of BCA-driven enterprises and limited data availability, the literature in this subsection mostly relies on surveys and theoretical concepts.

Research Gaps:

Although previous studies examine blockchain adoption frameworks, cryptocurrency sentiment, and blockchain applications in Vietnam separately, limited research investigates how blockchain-related sentiment may relate to broader economic indicators. This study addresses this gap by integrating sentiment analysis with macroeconomic variables.

- ***Considering Macroeconomic Indicators:*** Our paper occupies a unique position compared to prior work in frameworks and determinants. Rather than reinforcing BCA determinants stipulated by theoretical models such as TAM or TOE, our objective is to evaluate the early economic impact of BCA by examining its relationship with macroeconomic indicators, a distinct category of determinants that deviates from prior studies. Unlike Guo *et al.* and similar studies,²⁸⁻³² whose works are foundational to standard econometric models and most comparable to ours, our study engineers an innovative blockchain sentiment approach. Crucially, we encompass Vietnam's regulatory timeline to strengthen the rigor of BCA assessment rather than treating cryptocurrency adoption as the main focus.

- ***Extending the Application of News Sentiment Beyond Price Prediction:*** Existing papers implementing sentiment analysis predominantly focus on the interconnection between news and historical prices of cryptocurrencies. Extending its application beyond price prediction, we selectively leverage news sentiment as a measurable proxy, given the pivotal role of Vietnamese news media in interpreting the trajectory of BCA.

- ***Data-Driven Insights:*** Supply chain management is a particularly narrow topic of research in Vietnam BCA. Studies heavily rely on surveys due to the unavailability of data, leaving a vacuum of case studies on other blockchain applications. As we present more robust, data-driven economic insights through correlation, regression, and Granger causality, our assessment of BCA helps illuminate the nascent stage of BCA for future research to expand upon.

Taken together, this work bridges these gaps to deliver the first integrated economic perspective on BCA in Vietnam.

■ Methods

In this section, we elaborate on the steps taken to conduct the empirical research. First, we describe the data scraped from Vietnamese news articles, followed by the translation process. Second, we design techniques for sentiment analysis and article categorization. Next, we introduce both quantitative and qualitative metrics. While the former are aligned into time-series datasets and analyzed using Pearson correlation, regression, and Granger causality methods, the latter provide contextual observations. Finally, we briefly describe the coding environment used to implement the methodology.

Data Preprocessing and Sentiment Analysis:

Due to unstructured HTML layouts of Vietnamese news (e.g., formatting errors, duplicates of topics, incoherent configuration), there are no open-source datasets on articles with specific keywords targeting ‘blockchain’, ‘crypto’, or ‘bitcoin’. As a result, we manually create our own pipelines, using the BeautifulSoup library to extract thousands of articles by title, URL, date of publication, and content. We specifically hand-pick four domestic outlets based on their high reputation and credibility. The leading newspaper is VnExpress, characterized by fast and diverse news coverage and representing the public discourse of socioeconomic shifts in Vietnam.⁴⁹ The secondary popular choices are Tuoitre, Vietnamnet, and VnEconomy, but each has its own niche.⁵⁰⁻⁵² While Vietnamnet publishes the government’s voice on reforms and policies, revealing the roadblocks of BCA, Tuoitre offers a youth-focused perspective. VnEconomy’s specialization in economics and business helps contextualize the role of macroeconomic indicators in a broader context.

Facilitating international academic accessibility, we import the Google Translate module to translate article titles and contents into English. This enables our research to perform sentiment analysis with both English and Vietnamese pretrained NLP models, thereby normalizing the benchmark across models. The four models we used are ElKulako/cryptobert, ProsusAI/finbert, wonrax/phobert-base-vietnamese-sentiment, and mr4/phobert-base-vi-sentiment-analysis, all available through the Hugging Face repository.⁵³⁻⁵⁵ We justify the selection of models by their performance on past studies in the realm of financial text filled with blockchain and crypto lexicons. FinBERT is effective in this context, as it is fine-tuned from a large corpus of corporate reports and financial PhraseBank news. Its sentiment classification achieves an accuracy of 97%, outstripping the performance of Long Short-Term Memory (LSTM) models.⁵⁴ We also deploy CryptoBERT, a similarly advanced BERT model trained on cryptocurrency corpora, to capture nuances in crypto-specific terminology where generic financial interpretation may underperform.⁵³ For Vietnamese models, we implement two variants of PhoBERT models: Mr4 and Wonrax.⁵⁵ PhoBERT is a state-of-the-art model for Vietnamese-specific NLP. To ensure robustness, we inspect a small sample by randomly se-

lecting ten articles. All the NLP models achieve 90% accuracy in correctly identifying sentiment values.

The sample size of the paper is small. As manually reviewing and labeling Vietnamese blockchain-related articles is highly labor-intensive, the validation step was primarily designed as a robustness check rather than a full model training process. The sentiment models used in this study (i.e., FinBERT, CryptoBERT, and PhoBERT variants) are already pre-trained and validated by previous studies. Thus, the purpose of the validation was to evaluate whether these models could capture sentiment in Vietnamese blockchain news reasonably, instead of evaluating model performance from scratch.

After sentiment analysis, the dataset now contains two new columns: sentiment value [bullish, neutral, bearish], indicating the overall polarity of the text; and confidence score (0–1), indicating the model’s probability weighted for the assigned sentiment class. The dataset was then processed using a defined prompt on the OpenAI API platform to classify articles as ‘national’, ‘international’, and ‘irrelevant’. The batches of articles, originally separated by keywords and outlets, are combined and sorted chronologically from January 2018 to July 2025 (articles before 2018 are excluded). To prepare for time-series analysis, sentiment scores are aggregated by month to observe how sentiment changes. We compute that neutral articles contribute a score of 0, bullish articles contribute +1 multiplied by their confidence score, and bearish articles contribute -1 multiplied by their confidence score (Equation 1). The average sentiment score is then computed as the total aggregated sentiment divided by the number of articles.

$$AS = \frac{\sum_{i=1}^N S_i}{N}$$

where:

AS = Average Sentiment

S_i = Sentiment score of article i

N = Total number of articles in a given period

Equation 1: The equation, the average sentiment method, computes the overall polarity of all articles in a given period of time.

Quantitative Data:

Our granular quantitative dataset comprises eight variables: two sentiment indices derived from international and national news articles, and six macroeconomic indicators compiled from multiple sources. The macro-indicators are the exchange rate, inflation rate, foreign exchange reserve, balance of trade, foreign direct investment in blockchain-related sectors, and industrial production index.⁵⁶⁻⁶¹ Owing to reporting lags and data availability, the monthly time-series spans January 2020 to January 2025. Our explanation of these indicators is included below.

Exchange rate (EXR) denotes the rate at which a country’s currency can be exchanged for another.²⁸ In our study, EXR is expressed as USD/VND to reflect the monetary stability of the Vietnamese Dong (VND) because the dollar is the dominant global reserve currency. EXR, grounded in monetary policies

and economic theories, is a marker for digital activities, shaping economic development and BCA.^{28,30,31} When the dollar appreciates against VND (e.g., high EXR), VND diminishes its purchasing power, raising the cost of cross-border payments and remittances.³¹ Such volatility encourages individuals to adopt cryptocurrencies both as a store of value and a decentralized transaction medium that bypasses traditional high-cost channels.²⁹ While fluctuations in EXR can undermine financial stability, they stimulate capital flows into digital financial infrastructure and, by extension, broader BCA.

Inflation rate (INF) is the percentage increase in the price level of goods and services over time. Akin to EXR, higher INF strengthens the appeal of hedging assets like cryptocurrencies. Nguyen & Nguyen report that an inflation crisis often results in a higher crypto adoption rate,³² which is also consistent in Guo *et al.*'s analysis across 37 countries.²⁸

FX reserve (FXR) acts as the ultimate barometer of monetary confidence, yet it has been neglected by prior research on this subject. Largely dollar-denominated, it is a powerful intervention tool that the State Bank of Vietnam (SBV) could use to prevent the devaluation of VND.⁶² For an emerging country, where remittances and external trades are vitally integrated in the financial market, a high FX reserve signals a healthy economic outlook. It implies a stable exchange rate, a cushion against external shocks (e.g., trade imbalances or capital outflows), confidence in investors, and the ability to service international debt.^{63,64} Interestingly, Vietnam's FX reserves peaked in 2022. This hints at Vietnam's long-term strategy to achieve the "National Credit Rating Improvement Project by 2030" in tandem with its goal to become a central blockchain hub.⁶⁴ Especially, given that crypto transactions under the economy's shadow exceed 100 billion dollars,¹⁸ FX reserves—combined with EXR and INF—will present a deeper understanding of the linkage between monetary uncertainty and increases in digital asset investment.

Balance of trade (BoT) determines whether a country has a trade surplus or deficit; in other words, it is the difference between the amount of exports and imports of goods and services. Although its role in BCA is underexplored, Chang *et al.* stress how blockchain enhances transparency, trust, and efficiency in the supply chain.²² Therefore, BoT is a valid measurement for the use of blockchain solutions in trade.

Foreign Direct Investment (FDI) refers to the capital inflows into domestic businesses from foreign entities. Since Vietnam necessitates large-scale R&D to drive burgeoning digital infrastructure, incorporating FDI in the macroeconomic spectrum of BCA is essential.^{30,32} We define FDI blockchain-related sectors (FDI-BS) to be composed of three elements: professional, scientific, and technical activities; information and communication; and financial, banking, and insurance activities. Our rationale is that technology transfer, expertise, international alliances, and financial support can cultivate interest in fintech startups and digital trade platforms. A strong FDI-BS indicates growth in the digital ecosystem, and a low FDI-BS, conversely, might suggest weak governance and regulatory barriers.

The Industrial Production Index (IPI) is our last macroeconomic indicator, measuring the production output of the manufacturing and mining sectors. IPI is critical because Vietnam's economy is manufacturing- and export-oriented (e.g., processed seafood, electronic components, metal products). In general, it indicates how technologically advanced the overall economic infrastructure is. Also, its impact mirrors that of the supply chain in BoT. Growth in IPI could reflect the integration of blockchain to streamline the manufacturing process.

Certain indicators, such as GDP, unemployment, and remittance inflows, are deliberately excluded from our methodology for several reasons. In Vietnam, GDP is public service-driven, whereas the BCA only accounts for a marginal share of overall economic activity. Similarly, we speculate that structural unemployment, a primary form of unemployment in wholesale and retail sectors, also bears little relevance to BCA. Remittance inflow, in theory, can explain how monetary policies could influence payment activities. However, as Vietnam's financial governance remains non-digitalized, signs of substitution toward blockchain-based channels are nearly undetectable.

The omission of the nominal interest rate (INT) is due to a lack of data. INT interacts with monetary policy to a greater extent. It refers to an additional proportion, by annual percentage, that borrowers repay lenders on debt. Higher INT likely tightens up liquidity and discourages investment. Conversely, lower INT makes the cost of borrowing cheaper and depreciates VND, sparking investments. Similar to EXR, it can be associated with higher BCA.²⁸ That said, INT does not guarantee FDI because investors also expect long-term regulatory certainty and market optimism, which cannot be offset by favorable lending conditions alone. But that raises a problem—INT is sticky and only varies by year. This fixation produces lags in data updates, which is incoherent with our monthly timeframe and makes it elusive to obtain data (e.g., INT is only scaled to the year 2023 on TradingEconomics).

Qualitative Data:

Moving on to qualitative data, we collect policy context from government sources, digital adoption rate (DAR), and Network Readiness Index (NRI) from annual reports by Chainalysis and Portulans Institute, respectively.^{18,65-73} We utilize qualitative data to interpret findings in quantitative data.

Policy context is constructed to support the reliability of news sentiment and macroeconomic indicators. It is established as a solid source of data—a concrete timeline (2018-2025) of government BC-related decisions, regulations, and initiatives—that narrates the outcome of BCA. The voice of government adds depth for correlation and causation, hedging against mere speculation from other dimensions. The progression in policies indicates a positive government stance on BCA, stemming from skepticism to endorsement of BC technology.

Digital adoption rate (DAR), collected from 2020 to 2024, measures the population's rate of digital adoption, such as the internet, mobile payments, and digital wallets. **Network Readiness Index (NRI)**, collected from 2022 to 2024, indicates the impact of internet and communication technology (ICT) by

skill and widespread infrastructure levels. In Vietnam, ICT lies in the strength of e-commerce legislation, internet traffic and subscriptions, mobile apps development, and R&D of emerging technologies.⁶⁷ Vietnam also ranks fifth globally in adoption rate, especially in cryptocurrency trading.¹⁸ However, it is important to note that this reflects a limited form of BCA, as broader benefits of BCT, such as DeFi applications, supply chain viability, and digital governance, have yet to be fully leveraged.

Regression Analysis:

$$Y_i = \beta_0 + \beta_1 \text{nat_sen}_{1i} + \epsilon_i \text{ (Model-1)}$$

$$Y_i = \beta_2 + \beta_3 \text{int_sen}_{1i} + \epsilon_i \text{ (Model-2)}$$

Where Y_i is the dependent variable for observation i , referring to the six macroeconomic indicators. β_0 and β_2 are the constant terms, representing the expected value of the dependent variable when the independent variable equals zero. β_1 and β_3 are the coefficients of the independent variables, representing national and international sentiment in Models 1 and 2, respectively. ϵ_i is the error term, representing the difference between the actual value and the predicted value generated by the model. The estimated coefficients indicate how much the dependent variable changes when the corresponding independent variable increases by one unit.

Granger Causality Analysis:

Granger causality analysis is a statistical technique used to determine whether one time-series variable contains predictive information about another variable. The general Granger causality framework is based on the Vector Autoregression (VAR) model, in which each variable is expressed as a function of its own lagged values and the lagged values of other variables.

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \alpha_3 \text{nat_sen}_{t-1} + \alpha_4 \text{nat_sen}_{t-2} + \epsilon_t \text{ (Model-3a)}$$

$$\text{nat_sen}_t = \alpha_5 + \alpha_6 \text{nat_sen}_{t-1} + \alpha_7 \text{nat_sen}_{t-2} + \alpha_8 Y_{t-1} + \alpha_9 Y_{t-2} + \epsilon_t \text{ (Model-3b)}$$

$$Y_t = \alpha_{10} + \alpha_{11} Y_{t-1} + \alpha_{12} Y_{t-2} + \alpha_{13} \text{int_sen}_{t-1} + \alpha_{14} \text{int_sen}_{t-2} + \epsilon_t \text{ (Model-4a)}$$

$$\text{int_sen}_t = \alpha_{15} + \alpha_{16} \text{int_sen}_{t-1} + \alpha_{17} \text{int_sen}_{t-2} + \alpha_{18} Y_{t-1} + \alpha_{19} Y_{t-2} + \epsilon_t \text{ (Model-4b)}$$

Where Y_t represents the dependent variable at time t , while nat_sen and int_sen represent the explanatory variables. ϵ_t denotes the error term. The null hypothesis of the Granger causality test states that the lagged values of sentiment variables do not jointly affect Y_t . If the null hypothesis is rejected, then the sentiment variable Granger-causes Y_t , meaning that historical information from sentiment contributes significantly to forecasting Y_t . If both variables Granger-cause each other, a bidirectional relationship exists. Conversely, if neither variable Granger-causes the other, then no predictive relationship exists between them. In this paper, the Granger causality method is employed to examine whether changes in media sentiment contain predictive information regarding macroeconomic indicators and whether macroeconomic indicators can similarly predict future movements in sentiment.

Coding Environment:

Our work is demonstrated through our GitHub repository, which includes full pipelines from data scraping and preprocessing to statistical and econometric analyses.⁷⁴ The coding scripts were implemented in the Python 3.12 environment on VSCode. Core libraries are BeautifulSoup for scraping, Pandas for data structure, Scikit-learn and Transformer modules for NLP models, and Matplotlib for plotting and visualization. We have documented the repository so that it is comprehensible and replicable in minimal steps.

Findings

After applying the methodology presented above, the final dataset consists of 3546 blockchain, 4721 bitcoin, and 478 crypto articles. We notice that 'crypto' articles are drastically low, solely reported from a single outlet—Vietnamnet. Preliminary trials underline that articles filtered by the latter two keywords are prone to cyclic speculation and high-frequency swings, adding noise to our analysis. Blockchain articles show stable sentiment scores, thus constituting our primary dataset. Our findings are divided into three subsets: (a) Performance of NLP models; (b) Sentiment trends and macroeconomic variables; (c) Statistical Relationship between Sentiment and Economic Indicators.

Mr4 and Wonrax models will capture sentiment in national articles of original Vietnamese text; FinBERT and CryptoBERT score sentiment in international articles, both before and after translation. Since presenting all models would be redundant and complex, we rely on FinBERT and Mr4 figures for (b) and (c). Due to data constraints discussed in the methodology section, the period of sentiment is trimmed from January 2020 to January 2025 to align with economic factors. To preserve the integrity of sentiment collection, subset (a) was conducted on the full timeline and interpretation of international models across all categories within blockchain articles.

Performance of NLP models:

The subsection showcases the benchmark between CryptoBERT and FinBERT, and a distribution of sentiment classified by each NLP model. It validates our decision to examine FinBERT analysis over the others and outlines an overall outlook of sentiment across all articles. Figure 1 displays the effect of translation on sentiment scores, strictly when the sentiment value of an article is identical. After translation, CryptoBERT shows a slight positive bias (+0.0365), in which the frequencies of the delta score cluster around zero. FinBERT exhibits a negative bias (-0.0950). Although its deviation leans to the left, most of the articles are centered around the interval (-0.1 to 0.1). The translation effect reflects FinBERT's higher sensitivity towards linguistic nuances; nevertheless, the difference between the two models is minimal.

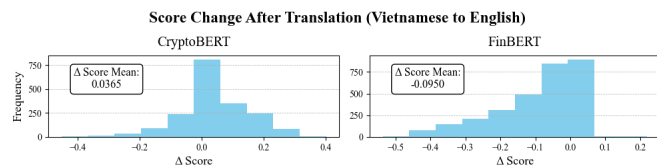


Figure 1: Histogram of sentiment score changes before and after translation for CryptoBERT and FinBERT. While CryptoBERT shows a slight positive bias (+0.0365), FinBERT exhibits a negative bias (-0.0950) after translation.

Figure 2, which compares sentiment matches, reveals a stark contrast. Both models show that the percentage of matching exceeds that of mismatching. However, CryptoBERT shows an even split, retaining 52.8% of articles, whereas FinBERT achieved a much higher proportion of matches, up to 84%. The figure suggests that FinBERT is a stronger candidate for interpreting international articles because its reliability score is superior to that of its counterpart.

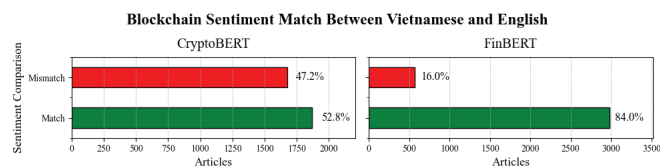


Figure 2: Comparison of sentiment matches and mismatches before and after translation for CryptoBERT and FinBERT. FinBERT achieves a substantially higher match rate (84.0%) than CryptoBERT (52.8%), suggesting greater reliability for sentiment analysis of translated articles.

In Figure 3, the sentiment is unequally distributed among the four models tested. Neutral and bullish values stand out as an eminent representation of articles, implying the impartiality of the writing tone and a glimpse of buoyancy in BC. Still, the partition of sentiment values deduced by each model differs substantially. Before translation, bearish sentiment is infinitesimal, with CryptoBERT detecting 0.3% and FinBERT detecting none at all. FinBERT insists that the proportion of bullish articles, which is relatively rare, sits at 0.4%, mirroring that of the bearish class. CryptoBERT's finding is opposite, identifying 34.1% of articles to be positive. This disparity leads to a dissimilarity in neutral classification: whereas CryptoBERT assigns 65.6% of articles to this value, FinBERT assigns 93.6% of them. After translation, both models agree that the number of negative articles was underestimated. CryptoBERT and FinBERT, respectively, re-evaluate negative articles to make up 4.1% and 8.0% of the dataset, though pessimism coverage is seemingly negligible. CryptoBERT shifts previously neutral articles into bullish, highlighting a transition in sentiment dominance. Such a shift also occurs in FinBERT's analysis; however, neutral articles remain the majority.

Classification by both Vietnamese models paints an overarching tone of bullishness. Mr4 registers 56.7% of articles as bullish, 18.6% as neutral, and 24.6% as bearish. Wonrax corresponds with the trend. It treats 68.6% of articles as bullish—a slightly higher proportion—15.8% neutral, and 24.6% bearish—a slightly lower proportion.

The results suggest that both PhoBERT variants are better attuned to local nuances because they specialize in the Vietnamese language. For this reason, we interpret that the

media tone of blockchain is confidence-skewed and distinctly forward-looking. In contrast, international models show that there is a low polarity of sentiment because articles are interpreted as overwhelmingly neutral. FinBERT clearly portrays this pattern, insinuating a cautious stance on the blockchain narrative in Vietnam.

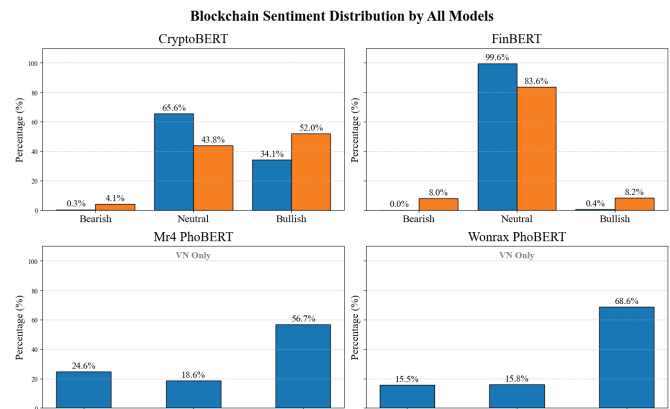


Figure 3: Percentage distribution of bearish, neutral, and bullish sentiment classifications across all NLP models. The results show that both PhoBERT variants are better attuned to capturing Vietnamese-written articles, as they are trained specifically for this language.

Sentiment Trends and Macroeconomic Variables:

For this subset, we present a trimmed timeline of how the sentiment score, captured by Mr4 PhoBERT, and all variables, before being embedded in the correlation map, progress by month.

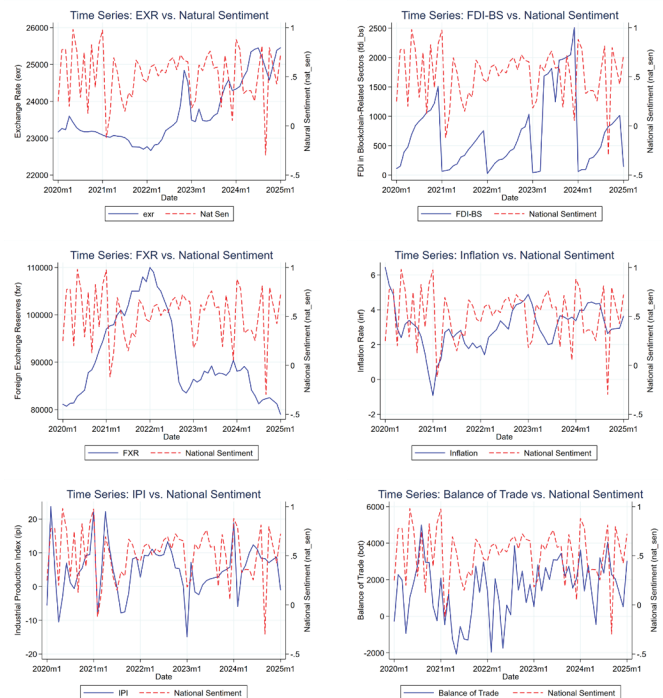


Figure 4: Time-Series Trends of National Sentiment and Macroeconomic Indicators in Vietnam (2020–2025). Visual patterns suggest a possible relationship between national sentiment and industrial production and trade activities; however, statistical analyses are required to confirm these relationships.

Figure 4 shows the relationship between national sentiment and the six macroeconomic indicators over time. Over the five years, national sentiment maintains a generally positive outlook, with fluctuations ranging between approximately 0.25 and 0.75 and a monthly average of 0.52. During the first year, sentiment shows an upward movement toward values between 0.8 and 1.0. Although national sentiment experienced a slight decline below 0.5 in mid-2023, it remained above 0.5 during most of the subsequent periods. Overall, despite several fluctuations, national sentiment demonstrates resilience and follows a more positive long-term trend toward 2025.

The macroeconomic indicators display different patterns alongside national sentiment. EXR steadily increases from around 23,000 to 25,500 USD/VND, with a noticeable rise between 2022 and 2023. FXR reached its highest level of over 110 billion USD during the first two years before declining toward 2025. IPI shows substantial fluctuations, repeatedly moving between -15% and +20%. Similarly, BoT alternates between trade surpluses and deficits, although surpluses become more dominant after 2021. Inflation decreases from approximately 6% to below 0% in 2021 and later stabilizes between 3% and 4%. FDI-BS generally shows an increasing trend over time despite a temporary decline in mid-2023. Overall, visual patterns suggest a possible relationship between national sentiment and industrial production and trade activities; however, statistical analyses are required to confirm these relationships.

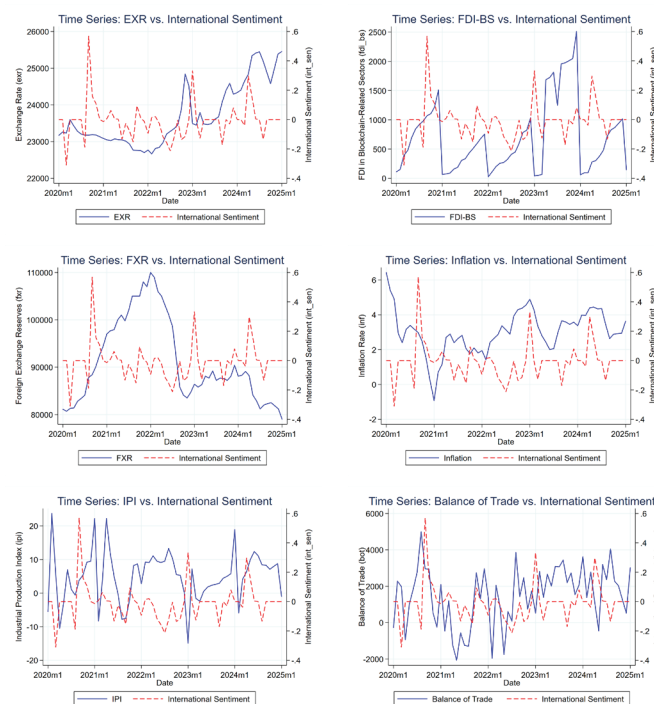


Figure 5: Time-Series Trends of International Sentiment and Macroeconomic Indicators in Vietnam (2020–2025). The relationships between international sentiment and macroeconomic indicators appear weaker than those observed for national sentiment.

Figure 5 illustrates the time series relationship between international sentiment and the six macroeconomic indicators. International sentiment has a lower monthly mean of 0.11, and it tends to stay closer to neutral values during the study period.

International sentiment was negative for a number of quarters in a row in the first year, then improved at the end of 2020, and then slowly declined. International sentiment dropped in mid-2023, but recovered quickly. Except for these fluctuations, the overall direction of international sentiment is slowly upwards to 2025.

While macroeconomic indicators still exhibit long-term trends and short-term fluctuations, international sentiment does not appear to follow these variables closely. International sentiment is less correlated with Vietnam's economic conditions than national sentiment, indicating that it may be indicative of more global conversations and external perspectives rather than domestic economic developments.

Statistical Relationship between Sentiment and Economic Indicators:

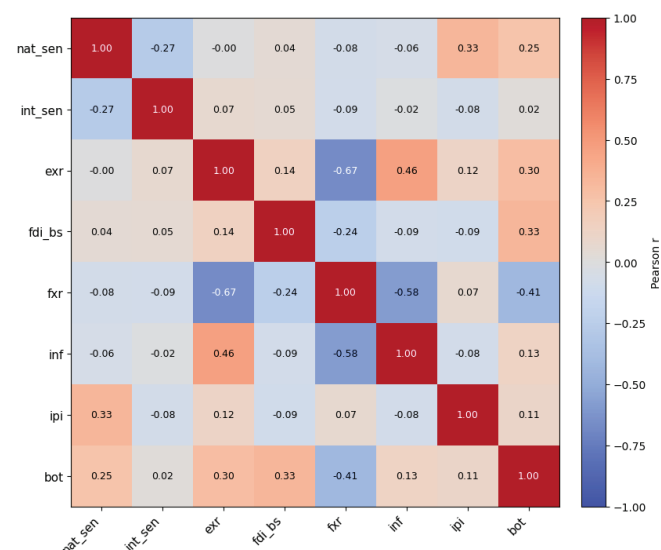


Figure 6: Heatmap of Pearson correlations between national sentiment (Mr4 PhoBERT), international sentiment (FinBERT), and six macroeconomic indicators from January 2020 to January 2025. The score ranges from -1 (perfectly negatively related) to +1 (perfectly positively related), and 0 shows non-correlation. National sentiment exhibits relatively stronger positive correlations with IPI (0.33) and BoT (0.25), while both sentiment variables show weak correlations with most other macroeconomic indicators.

As shown in Figure 6, both sentiment variables exhibit weak or negligible correlations with most macroeconomic indicators, with coefficients ranging between -0.09 and 0.07. National sentiment shows relatively stronger positive correlations with IPI (0.33) and BoT (0.25). Stronger relationships also emerge within the macroeconomic indicators themselves. Positive correlations include EXR-INF (0.46), EXR-BoT (0.30), and FDI-BS-BoT (0.33). Negative correlations include EXR-FXR (-0.67), FXR-BoT (-0.41), and FXR-INF (-0.58). Given the focus of the paper on BCA in Vietnam, the discussion primarily interprets the relationship between national sentiment, IPI, and BoT, which is supported by qualitative observations.

Nat_sen and IPI: National sentiment (-0.2 to +0.9) and IPI (-10% to +22%) show the widest fluctuations in their first year interval. At the same time, Vietnam's DAR ranked 10th globally: On-chain value and retail value received ranked the second highest, while the numbers of on-chain deposits (44th)

and P2P exchange trade volume (81st) remain low. This discrepancy among the sub-indices signals high value flows due to transfer activities, but the grassroots in exchange and wallet engagement are not yet mainstream.

The May 2020 directive from the Prime Minister instructing the SBV to study blockchain utility suggests the exploratory and cautious state of adoption. However, as DAR reaches the first place in the following years amid surges in P2P trade volume and DeFi protocols, national sentiment (-0.2 to +0.7) and IPI (-10% to +12%) both rebound and steadily increase.

In the NRI context, Vietnam scores well above its corresponding GDP per capita. Strong sub-pillar, including individuals and the economy, helps Vietnam outperform in digital readiness compared to similar-income-level countries. Ongoing policies toward consumer protection may have fueled this development. A legal entity for blockchain established in March 2022 sets a new milestone for adoption.

The grassroots technology infrastructure continues to expand, accompanied by improvements in NRI ranking from 62/131 countries in 2022 to 45/133 in 2024. While DeFi and centralized exchanges remain active, DAR declines to third and fifth place. During this period, there is a broader decline in P2P trading and a hiatus in BC policies.

The drafting of the crypto framework in February 2024 and Resolution No. 57-NQ/TW, passed later that year, coincide with temporary dips in national sentiment and IPI at the beginning of both years, followed by upward trends later in the year. This dynamic shows an annual cycle where confidence gradually reduces early uncertainty through improvements in NRI performance and renewed policy support. This cycle also supports the moderate relationship between nat_sen and IPI.

Nat_sen and BoT: The graph of BoT exhibits a pattern similar to IPI, and certain outliers and trends align with movements in national sentiment. There were alternating surpluses and deficits across 2020, an era where the world experienced a pandemic. The COVID-19 lockdown precipitated major bottlenecks in logistics and hindered production rate, thereby supporting the extensive volatility in both exports and imports. Media tone also reflects the downturn, where the rhythm of low optimism (below 0.2) preceding high optimism (0.9) took place throughout that year. Blockchain was yet to be an anchor because the adoption status was high value but low diffusion, given the uneven statistics of the 2020 DAR. In early 2021, national sentiment dropped below neutrality, and BoT recorded a deficit of 2 billion USD. P2P trade volume soars dramatically. However, it appears as a hedge against fluctuations in monetary tenets (e.g., currency devaluation and higher inflation rates), rather than a cue for rejuvenating the economy. The absence of a clear regulatory framework evokes uncertainty. It was not until the announcement that the BC association promised to nurture the cryptocurrency environment in March 2022 that imports exceeded exports by nearly 4 billion USD in mid-2022. This is parallel to the rise of bullish sentiment in media coverage, performing consistently from 0.5 to 0.8.

National sentiment falls short in late 2023 and quarter 3 of 2024. BoT is observed to fall short in quarter 1 and late 2024. This lag implies that shifts in sentiment may influence trade

performance. Neither the policy timeline nor NRI explains this delayed correlation. However, we hypothesize that it stems from weakening demand in export markets of Vietnam, rising costs of imports, or depreciation of the dong, which would make Vietnamese exports cheaper relative to foreign goods.

Although Pearson correlation is useful for drawing qualitative observations, it only provides linear relationships and does not account for statistical significance or predictive interactions. Therefore, to further investigate whether these relationships remain after controlling for statistical effects, regression analysis is conducted. The results are presented in Tables 1 and 2 below.

Table 1: Regression Analysis for National Sentiment (nat_sen). National sentiment has a significant positive relationship with IPI and BoT, while no significant associations are observed for EXR, FDI-BS, FXR, and INF.

VARIABLES	(1) EXR	(2) FDI-BS	(3) FXR	(4) INF	(5) IPI	(6) BoT
nat_sen	-102.471 (388.611)	76.550 (324.767)	-2,315.285 (3,793.814)	-0.337 (0.770)	9.460** (4.293)	1,502.355* (752.143)
Constant	23,729.296*** (237.349)	649.182*** (198.285)	92,473.952*** (2,333.646)	3.213*** (0.420)	0.099 (2.372)	701.928 (448.583)
Observations	60	60	60	60	60	60
R-squared	0.001	0.001	0.005	0.005	0.112	0.062

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 1 reports the results of the regression analysis of the association between the national media sentiment (nat_sen) (independent variable) and six macroeconomic variables (dependent variable): exchange rate (EXR), foreign direct investment in blockchain-related sectors (FDI-BS), foreign exchange reserves (FXR), inflation rate (INF), industrial production index (IPI), and balance of trade (BoT).

The coefficients provide us with an idea about the direction and magnitude of the association between the sentiment and the six macroeconomic factors. A positive coefficient suggests that an increase in national sentiment is associated with an increase in the relevant macroeconomic indicator, while a negative coefficient suggests the opposite. Robust standard errors are used to evaluate statistical significance at the 1% (***), 5% (**), and 10% (*) levels.

The evidence shows that national sentiment (nat_sen) has a statistically significant and positive impact on the Industrial Production Index (IPI) at a 5% significance level ($\beta = 9.460$, $p < 0.05$). The results indicate an improved attitude in national media and an increase in industry activity and output. This shows that higher national media sentiment is connected with increased industrial activity and productivity. One possible reason could be that favorable news and opinions may boost trust in technological progress and economic development, which may help sustain industrial expansion.

A similar finding is obtained for the Balance of Trade (BoT), where national sentiment is positively related at a 10% significant level ($\beta = 1,502.355$, $p < 0.10$). This conclusion implies that an increase in domestic sentiment may help achieve better trade performance, possibly through the channel of improved confidence in economic activity and technical progress. The data do not support any statistically significant correlations between national sentiment, exchange rates, blockchain-related foreign direct investment, foreign currency reserves, or inflation.

These findings indicate that the direct effect of media attitude on monetary and financial variables may be limited, as such variables are more likely to be influenced by wider macroeconomic conditions and policy issues.

Table 2: Regression Analysis for International Sentiment (int_sen). None of the estimated coefficients is statistically significant. Rather, there is a weak association between international media sentiment and economic conditions in Vietnam.

VARIABLES	(1) EXR	(2) FDI-BS	(3) FXR	(4) INF	(5) IPI	(6) BOT
int_sen	423.962 (740.529)	246.676 (481.235)	-6,053.806 (6,154.353)	-0.085 (1.097)	-4.908 (7.340)	179.070 (1,448.299)
Constant	23,679.825*** (108.515)	690.619*** (81.848)	91,231.860*** (1,139.451)	3.039*** (0.164)	4.938*** (0.998)	1,478.091*** (208.709)
Observations	60	60	60	60	60	60
R-squared	0.004	0.003	0.008	0.000	0.007	0.000

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 2 reports the regression results of the relationship between international media sentiment (int_sen) (explanatory variable) and six macroeconomic indices (dependent variable). The results indicate that none of the estimated coefficients is statistically significant. Rather, there is a weak association between international media sentiment and economic conditions in Vietnam.

One reason is that international media may focus on global issues and external perspectives, which may not have a direct impact on Vietnam's domestic economy. The local media sentiment, on the other hand, could better capture country-specific developments and public perceptions regarding Vietnam's economy.

Regression analysis is helpful to find associations between sentiment and macroeconomic variables, but it does not tell us whether one variable can predict changes in another variable in the future. Therefore, Granger causality analysis is applied to study the predictive relationships between these variables.

Table 3: Granger Causality Test between Sentiment Variables and Macroeconomic Indicators. There is a statistically significant relationship between national sentiment and investment, trade, and inflation activities, and between international sentiment and broader financial and production indicators.

Sentiment	Equation	Excluded	chi2	Prob>chi2
National	exr	nat_sen	0.039	0.981
	nat_sen	exr	3941	0.139
	fdi_bs	nat_sen	13.769	0.017**
	nat_sen	fdi_bs	4.215	0.519
	fxr	nat_sen	0.8934	0.981
	nat_sen	fxr	3.2374	0.198
	inf	nat_sen	10.609	0.031**
	nat_sen	inf	2.2869	0.580
	ipi	nat_sen	4.413	0.110
	nat_sen	ipi	3.513	0.173
International	bot	nat_sen	12.323	0.055*
	nat_sen	bot	12.212	0.057*
	exr	int_sen	2.1887	0.728
	int_sen	exr	9.8956	0.078*
	fdi_bs	int_sen	0.7812	0.677
	int_sen	fdi_bs	0.9871	0.610
	fxr	int_sen	1.2755	0.735
	int_sen	fxr	12.254	0.007***
	inf	int_sen	13.939	0.007***
	int_sen	inf	3.4617	0.4842
	ipi	int_sen	3.5884	0.058*
	int_sen	ipi	3.8984	0.048**
	bot	int_sen	0.5509	0.759
	int_sen	bot	5.3106	0.070*

*** p<0.01, ** p<0.05, * p<0.1

Table 3 presents the results of the Granger causality test between sentiment variables and macroeconomic indicators. We use the Granger causality test to check whether lagged sentiment is able to forecast changes in macroeconomic indicators, and vice versa. The results indicate that for national sentiment (nat_sen), FDI-BS ($\chi^2 = 13.769$, $p < 0.05$), inflation ($\chi^2 = 10.609$, $p < 0.05$), and balance of trade ($\chi^2 = 12.323$, $p < 0.10$) significantly Granger-cause national sentiment. This suggests that past changes in these economic variables may be useful for predicting future changes in national sentiment. However, exchange rate, foreign exchange reserves, and industrial production are not strong predictors of national sentiment. It also finds a bi-directional relationship between trade balance and national sentiment. This suggests that trade performance affects sentiment, and sentiment may be helpful in predicting future trade performance. The results underscore the tight link between national sentiment and domestic economic conditions, especially investment, trade, and inflation activities.

The exchange rate is an important predictor of international sentiment (int_sen) at the 10% level ($\chi^2 = 9.896$, $p < 0.10$). Foreign exchange reserves and inflation significantly predict international sentiment at a 1% significance level. Besides, industrial production is also a predictor of international sentiment at the 10% level. The results also indicate that international sentiment is an important predictor of foreign exchange reserves, industrial production, and balance of trade. The bi-directional link between international sentiment and industrial production points to the possibility of an interaction between the variables across time.

In conclusion, the results of the Granger causality tests show that the sentiment and the macroeconomic indicators are related over time and not in a simple one-way causality. Domestic economic activity is more correlated with national sentiment, and broader financial and production indicators are more correlated with international sentiment.

■ Discussion

Limitations:

The methodology has inherent limitations that may potentially undermine the reliability of the findings. Collecting articles is challenging. Different media outlets have different HTML structures, and scraping features, such as APIs, are not readily accessible. The dataset also poses caveats. The titles from articles are varied, but sources can publish overlapping content, necessitating an intensive manual check. Biases in translation, sentiment scoring, and categorization systems by libraries are inevitable. Because the timelines of indicators are mostly disjointed, the time series is restricted to five years. Given the exploratory nature of the research, our rationale for correlations can be subjective and does not purport to be causations. More sophisticated econometric techniques could be implemented in future research to enhance the validity of the correlations. Last but not least, the lack of data-driven infrastructure in Vietnam continues to constrain empirical studies on the early impact of BCA.

Future Directions:

The research frontier can be expanded to address data issues as mentioned earlier. Broadening blockchain-related keywords and extending the timeline beyond the current span enables the enrichment of the dataset. Advanced methodologies such as the Autoregressive Distributed Lag (ARDL) model, machine learning techniques, and the Vector Error Correction Model (VECM) can introduce temporal lag structures, long-run equilibrium relationships, and predictive or non-linear interactions between sentiment and macroeconomic indicators. Comparative studies across ASEAN economies and sectoral analyses could offer distinctive research angles, further strengthen the findings, and comprehensively contextualize BCA in Vietnam.

Conclusion

In emerging markets, scalability and interoperability are essential to the rampant maturation of financial systems, monetary fundamentals, domain exchanges, and trustless transactions. The adoption of blockchain in Vietnam is no longer a discrete phenomenon—it is increasingly interwoven across various economic sectors, potentially reshaping conventional business infrastructures. Our paper calls attention to the emergence of blockchain and its economic implications by examining the relationship between media sentiment and macroeconomic indicators. To conclude the paper as a whole, we now attempt to answer the three guiding questions from the introduction based on our findings.

- RQ1: The paper measures sentiment by using NLP. There are notable correlations between news sentiment and macroeconomic indicators.

- RQ2: The findings suggest that shifts in Vietnamese media sentiment have a limited but noticeable relationship with blockchain-related economic activity. Pearson correlation, regression, and Granger causality yield relatively similar results, which could be plausibly supported by qualitative data to some extent.

- RQ3: The results indicate that the industrial production index (IPI) and balance of trade (BoT) exhibit significant relationships with national sentiment. In contrast, international sentiment does not show statistically significant relationships with the macroeconomic indicators examined.

All in all, our research is innovative and significant, yet proven. It contributes a valuable baseline to assessing the economic impact of blockchain adoption. It offers a synthesis of quantitative and qualitative datasets, redressing the fragmentation of existing research on this emerging technology and laying the ground for future work.

Acknowledgments

I would like to express my gratitude to Prof. Faisal Nawab and Prof. Abdullah Yalaman for their invaluable guidance and support throughout my research endeavors.

References

1. World Bank. (2021). *The Global Findex Database 2021*. Washington, DC: World Bank.

2. Doarest, A., & Wihardja, M. M. (2024). *The impacts of supply chain reconfiguration on ASEAN economies*. ISEAS–Yusof Ishak Institute.
3. International Federation of Accountants. (2023). *SMEs: Backbone of Southeast Asia's growing economy*. New York, NY: IFAC.
4. Ishaya, B. (2025). Problematizing socially sustainable global supply chains. *International Journal of Development Planning Literature*, 39(2), 101–120.
5. Suranto, B. (2025). State of digitalization in the Southeast Asia region. *Quality & Quantity: International Journal of Methodology*, 59(4), 1123–1145.
6. Tripathi, G., Ahad, M. A., & Casalino, G. (2023). *A comprehensive review of blockchain technology: Underlying principles and historical background with future challenges*. *Decision Analytics Journal*, 9, Article 100344.
7. Al-Sulami, Z. A., Ali, N., Ramli, R., & Lu, S. (2024). Towards a comprehensive understanding of blockchain technology adoption in various industries in developing and emerging economies: A systematic review. *Cogent Business & Management*, 11(1), Article 2294875.
8. Staley, I. (2025). Bridging blockchain and digital asset gaps: A comparative policy analysis of regulatory practices in emerging markets. *IET Blockchain*, 1(1), 45–62.
9. Schar, F. (2021). Decentralized finance: On blockchain- and smart contract-based financial markets. *Federal Reserve Bank of St. Louis Review*, 103(2), 153–174.
10. Peter Yeoh (2017). Regulatory issues in blockchain technology. *Journal of Financial Regulation and Compliance*, 8 May 2017; 25 (2): 196–208.
11. Michael, J., Cohn, A. L., & Butcher, J. R. (2018). Blockchain technology and regulatory investigations. *The Journal*, 1(7), 1–11.
12. Monetary Authority of Singapore. (2024). *Regulatory sandbox*. Monetary Authority of Singapore.
13. Baker McKenzie. (2025). *Philippines: SEC issues rules and guidelines on crypto asset service providers (CASPs)*. Insight+.
14. Giang, D. T. (2025). *Digitalization and its discontents: Paper-based public services in Vietnam*. *Journal of Public Administration and Governance*, 15(2), 87–105.
15. Chau, N. T. M., & Son, T. H. (2019). *Understanding financial inclusion in Vietnam*. *Review of World Economics*, 10(3), 382–390.
16. Nguyen, V. T., Le, T. H., & Phan, Q. P. (2025). *Traceability gaps in Vietnam's agricultural supply chains: Risks and remedies*. *Journal of Supply Chain Management*, 7(1), 45–60.
17. Tran, V. Q., & Nguyen, H. T. (2025). *Bureaucratic bottlenecks and productivity in Vietnamese provincial administrations*. *Asian Journal of Public Sector Governance*, 3(1), 23–41.
18. Chainalysis Team. (2024, September 11). *The 2024 global adoption index: Central & Southern Asia and Oceania (CSAO) region leads the world in terms of global cryptocurrency adoption*. Chainalysis. <https://www.chainalysis.com/blog/2024-global-crypto-adoption-index/>
19. VietnamPlus. (2025). *Vietnam urgently needs a legal framework for digital assets: Experts*. VietnamPlus.
20. LuatVietnam News. (2024). *Nation to become a leader in blockchain technology*. LuatVietnam News.
21. AlShamsi, M., Al-Emran, M., & Shaalan, K. (2022). A systematic review on blockchain adoption. *Applied Sciences*, 12(9), 4245.
22. Chang, Y., Iakovou, E., & Shi, W. (2020). Blockchain in global supply chains and cross-border trade: a critical synthesis of the state-of-the-art, challenges and opportunities. *International Journal of Production Research*, 58(7), 2082–2099.
23. Taherdoost, H. (2022). A Critical Review of Blockchain Acceptance Models—Blockchain Technology Adoption Frameworks and Applications. *Computers*, 11(2), 24.

24. Lou, A. T., & Li, E. Y. (2017). Integrating innovation diffusion theory and the technology acceptance model: The adoption of blockchain technology from business managers' perspective.
25. Nuryyev, G., Wang, Y.-P., Achyldurdyeva, J., Jaw, B.-S., Yeh, Y.-S., Lin, H.-T., & Wu, L.-F. (2020). Blockchain Technology Adoption Behavior and Sustainability of the Business in Tourism and Hospitality SMEs: An Empirical Study. *Sustainability*, 12(3), 1256.
26. Malik, S., Chadhar, M., Vatanasakdakul, S., & Chetty, M. (2021). Factors Affecting the Organizational Adoption of Blockchain Technology: Extending the Technology–Organization–Environment (TOE) Framework in the Australian Context. *Sustainability*, 13(16), 9404.
27. Chittipaka, V., Kumar, S., Sivarajah, U., Bowden, J. L. H., & Baral, M. M. (2023). Blockchain Technology for Supply Chains operating in emerging markets: an empirical examination of the technology-organization-environment (TOE) framework. *Annals of Operations Research*, 327(1), 465-492.
28. Guo, Y., Yousef, E., & Naseer, M. M. (2025). Examining the drivers and economic and social impacts of cryptocurrency adoption. *FinTech*, 4(1), 5.
29. Sergio, I., & Wedemeier, J. (2025). Global surge: exploring cryptocurrency adoption with evidence from spatial models. *Financial Innovation*, 11(1), 96.
30. Stupak, Vyacheslav. "Economic Risk and Cryptocurrency: What Drives Global Digital Asset Adoption?" *Journal of Risk and Financial Management* 18, no. 8 (2025): 453.
31. Magazzino, C., Gattone, T., & Horkey, F. (2025). Economic and financial development as determinants of crypto adoption. *International Review of Financial Analysis*, 103, 104217.
32. Nguyen, L. T. M., & Nguyen, P. T. (2024). Determinants of cryptocurrency and decentralized finance adoption: A configurational exploration. *Technological Forecasting and Social Change*, 201, 123244.
33. Godbole, N., Srinivasaiiah, M., & Skiena, S. (2007). Large-scale sentiment analysis for news and blogs. *ICWSM*, 7(21), 219-222.
34. Shapiro, A. H., Sudhof, M., & Wilson, D. J. (2022). Measuring news sentiment. *Journal of Econometrics*, 228(2), 221-243.
35. Khedr, A. E., & Yaseen, N. (2017). Predicting stock market behavior using data mining techniques and news sentiment analysis. *International Journal of Intelligent Systems and Applications*, 9(7), 22.
36. Coulter, K. A. (2022). The impact of news media on Bitcoin prices: modeling data-driven discourses in the crypto-economy with natural language processing. *Royal Society Open Science*, 9(4), 220276.
37. Caviggioli, F., Lamberti, L., Landoni, P., & Meola, P. (2020). Technology adoption news and corporate reputation: Sentiment analysis about the introduction of Bitcoin. *Journal of Product & Brand Management*, 29(7), 877-897.
38. Roumeliotis, K. I., Tselikas, N. D., & Nasiopoulos, D. K. (2024). LLMs and NLP models in cryptocurrency sentiment analysis: A comparative classification study. *Big Data and Cognitive Computing*, 8(6), 63.
39. Vu, L. T., Pham, D. N., Kieu, H. T., & Pham, T. T. T. (2023). Sentiments extracted from news and stock market reactions in Vietnam. *International Journal of Financial Studies*, 11(3), 101.
40. Dang, T. T. (2019). Current situation of cryptocurrency in Vietnam. *The Journal of Business, Economics, and Environmental Studies*, 9(4), 29-34.
41. Lam, N. T. (2024). Cryptocurrency regulations: An international perspective and implications for Vietnam (Master's dissertation, University of the West of England, Bristol). *Faculty of Business and Law, Bristol Law School*.
42. Trung, M. D., Nghia, N. D. D., Thinh, N. Q., Tung, H. M., Van Long, D., Son, T. H., ... & Tho, Q. T. (2019). An overview of blockchain applications and their impact on the Vietnamese context. *Publishing House for Science and Technology*, 28-35.
43. Le Minh, H., Diep, N. N., Anh, N. H., & Linh, D. P. (2024). Blockchain technology characteristics in supply chain management: Evidence from Vietnamese enterprises. *International Journal of Economics, Business and Entrepreneurship Research*.
44. Vu, T. T., & Trinh, H. H. H. (2021). Blockchain technology for sustainable supply chains of agri-food in Vietnam: a SWOT analysis. *VNUHCM Journal of Economics-Law and Management*, 5(1), 1278-1289.
45. Nguyen Hoang, K., Nguyen Duong Phu, T., Nguyen Minh, T., Hoang Luong, H., Tran Dang, K., Huynh, K., ... & Vo Hong, K. (2023a, November). Blockchain as a Collaborative Technology-Case Studies in the Real Estate Sector in Vietnam. In *Pacific Rim Knowledge Acquisition Workshop* (pp. 111-123). Singapore: Springer Nature Singapore.
46. Pham, C. T., & Nguyet, T. T. T. (2023). Determinants of blockchain adoption in news media platforms: A perspective from the Vietnamese press industry. *Heliyon*, 9(1).
47. Nguyen, L. T., Nguyen, D. T., Ngoc, K. N. N., & Duc, D. T. V. (2023b). Blockchain adoption in logistics companies in Ho Chi Minh City, Vietnam. *Cogent Business & Management*, 10(2), 2216436.
48. Nguyen, N. T., Nguyen, A. T., To, H. T. N., & Le, T. T. H. (2024). Why are Vietnamese people susceptible to cryptocurrency Ponzi schemes? Findings from using the PLS-SEM approach. *Journal of Financial Crime*, 31(1), 158-173.
49. VnExpress. *Homepage*. Retrieved June 14, 2025, from <https://vn-express.net/>
50. Tuổi Trẻ News. *Homepage*. Retrieved June 14, 2025, from <https://tuoitre.vn/>
51. VietnamNet. *Homepage*. Retrieved June 14, 2025, from <https://vietnamnet.vn/>
52. VnEconomy. *Homepage*. Retrieved June 14, 2025, from <https://vneconomy.vn/>
53. Kulakowski, M., & Frasinca, F. (2023). Sentiment classification of cryptocurrency-related social media posts. *IEEE Intelligent Systems*, 38(4), 5-9.
54. Araci, D. (2019). FinBERT: Financial sentiment analysis with pre-trained language models. *arXiv preprint arXiv:1908.10063*.
55. Nguyen, D. Q., & Nguyen, A. T. (2020). PhoBERT: Pre-trained language models for Vietnamese. *arXiv preprint arXiv:2003.00744*.
56. Reserve Bank of Australia. *Statistical tables: Historical data – exchange rates*. Retrieved August 2, 2025, from <https://www.rba.gov.au/statistics/historical-data.html#exchange-rates>
57. Trading Economics. *Vietnam inflation (CPI)*. Retrieved August 2, 2025, from <https://tradingeconomics.com/vietnam/inflation-cpi>
58. Trading Economics. *Vietnam's foreign exchange reserves*. Retrieved August 2, 2025, from <https://tradingeconomics.com/vietnam/foreign-exchange-reserves>
59. Trading Economics. *Vietnam's balance of trade*. Retrieved August 2, 2025, from <https://tradingeconomics.com/vietnam/balance-of-trade>
60. Trading Economics. *Vietnam industrial production*. Retrieved August 2, 2025, from <https://tradingeconomics.com/vietnam/industrial-production>
61. Ministry of Planning and Investment of Vietnam. *Foreign investment statistics*. Retrieved August 2, 2025, from <https://www.mpi.gov.vn/en/Pages/solieudautunuocngoai.aspx>
62. Ahmed, R., Aizenman, J., Saadaoui, J., & Uddin, G. S. (2023). On the effectiveness of foreign exchange reserves during the 2021-22 US monetary tightening cycle. *Economics Letters*, 233, 111367.

63. Council on Foreign Relations. (2023, May 10). *The dollar: The world's reserve currency*.
64. Agribank. (2023, December 25). *Vietnam aims to increase foreign exchange reserves to 16 weeks of imports*.
65. Portulans Institute. (2022). *Vietnam country profile: Network readiness index 2022*. Network Readiness Index.
66. Portulans Institute. (2023). *Vietnam country profile: Network readiness index 2023*. Network Readiness Index.
67. Portulans Institute. (2024). *Vietnam country profile: Network readiness index 2024*. Network Readiness Index.
68. Chainalysis Team. (2021, October 14). The 2021 global crypto adoption index: Worldwide adoption jumps over 880% with P2P platforms driving cryptocurrency usage in emerging markets. Chainalysis. <https://www.chainalysis.com/blog/2021-global-crypto-adoption-index/>
69. Chainalysis Team. (2022, September 14). The 2022 global crypto adoption index: Emerging markets lead in grassroots adoption, China remains active despite the ban, and crypto fundamentals appear healthy. Chainalysis. <https://www.chainalysis.com/blog/2022-global-crypto-adoption-index/>
70. Chainalysis Team. (2023, September 12). *The 2023 global crypto adoption index: Central & Southern Asia are leading the way in grassroots crypto adoption*. Chainalysis. <https://www.chainalysis.com/blog/2023-global-crypto-adoption-index/>
71. Mills, L. (2025, June 11). Vietnam's crypto legislation opens the door to new business opportunities. Crypto for Innovation. <https://cryptoforinnovation.org/vietnams-crypto-legislation-opens-door-to-new-business-opportunities/>
72. Anh, C., & Lee, Y. (2024, May 16). How is the crypto tax in Vietnam? Tiger Research. Compiled by Felix, PANews. <https://www.chaincatcher.com/en/article/2124607>
73. R., E. (2025, July 28). Crypto regulations in Vietnam 2025. Coinpedia. <https://coinpedia.org/cryptocurrency-regulation/cryptocurrency-regulations-in-vietnam/>
74. Pham, M. (2025). Blockchain Research [Computer software]. GitHub. https://github.com/MichaelPham-HF/Blockchain_Research/tree/Regression-and-Granger-causality-test

■ Author

Michael Pham is a rising senior at The Hill School, PA. His academic interests delve into quantitative economics, data science, and mathematics. Through empirical research and computational analysis, his work explores interdisciplinary economics, specializing in blockchain technology and decentralized finance (DeFi). He has completed coursework in multivariable calculus, machine learning, neural networks, and cybersecurity, and is currently studying Higher Mathematics, Proof Seminar, and Advanced Economics. Outside the classroom, Michael pursues independent study in financial markets and quantitative finance, with interests in investment banking, portfolio management, and the application of computational methods to tackle global economic challenges.

Enhanced Combination Therapy of Aminoglycosides Against Resistant *Staphylococcus aureus*

Olivia Liu

Hong Kong International School, 1 Red Hill Rd, Tai Tam, Hong Kong; liuliviaa@gmail.com

ABSTRACT: *Staphylococcus aureus* is a fatal bacterial infection that commonly colonizes the skin and mucous membranes in humans. Each year, *S. aureus* is associated with an estimated 50,000 deaths, and this number is only rising with the increased prevalence of the pathogen's resistance against many antibiotics. In particular, the emergence of Methicillin-resistant *Staphylococcus aureus* (MRSA) has made it difficult to find suitable treatments for patients without the use of common clinical antibiotics. Combination therapy, which administers two or more antibiotics at once, is an effective approach to increase antibiotic efficacy against these resistant pathogens. This study investigates the efficacy of five antibiotics against the *S. aureus* Newman strain and finds that the Newman strain had significant bacterial resistance to kanamycin, chloramphenicol, and gentamicin. Contrarily, cefotaxime and erythromycin were highly effective even at lower doses. Eight combination therapy groups were tested with the more resistant antibiotics to find that the combination of gentamicin (10 µg/mL) and kanamycin (50 µg/mL) yielded the most effective results; the enhanced result of the two aminoglycosides working in combination led to a 91.18% reduction in bacterial growth compared to the untreated control group.

KEYWORDS: Medical and Health Sciences, Microbiology, *Staphylococcus aureus*, Antibiotic Combination Therapy, Aminoglycoside Synergy.

■ Introduction

Staphylococcus aureus is a bacterial human pathogen that colonizes the skin and mucous membranes, identifiable by its “grape-like” clusters and golden pigmentation at the microbiological level. This widespread bacterium infects approximately 20 to 50 individuals in a population of 100,000 annually, with an estimated 10 to 30 percent of these cases leading to fatalities.¹ In most healthy individuals, *S. aureus* can appear asymptotically on the skin and nasal cavities. In fact, a report by the European Centre for Disease Prevention and Control (ECDC) confirms that about 30 percent of the human population persistently carries *S. aureus* in their nasal mucous membranes without showing significant symptoms.² However, when the *S. aureus* bacterium gains access to internal organs or the bloodstream, it can cause severe infections, including sepsis and pneumonia.³ In the United States alone, over 119,000 cases of *S. aureus* entering the bloodstream were recorded in 2017, leading to an estimated 20,000 deaths from *S. aureus* just that year.⁴

Considering the nature of *S. aureus* as a bacterial pathogen, the treatment of *S. aureus* infections typically relies on antibiotic therapies, including commonly used agents such as cefazolin, nafcillin, oxacillin, and vancomycin.⁵ Yet, despite the prevalence of *S. aureus*, an antibiotic treatment specifically targeting all *S. aureus* infections is still elusive, particularly due to rising antibiotic resistance across strains of *S. aureus*. To test the efficacy of certain antibiotics on a human strain of *S. aureus*, we used five different antibiotics in varying concentrations. Among the antibiotic treatments we explored in this study, kanamycin is an aminoglycoside antibiotic with resistance rates exceeding 90 percent in many strains of *S. aureus*. Its mechanism of action

involves its irreversible binding to the 16S rRNA receptor in the bacterial 30S ribosomal subunit, causing a change in the translation process.⁵ Because of the faulty readings of codon and anticodon pairing, the resultant polypeptide becomes deformed, and the accumulation of these abnormal proteins inside the bacteria causes the cell to undergo apoptosis due to its high toxicity.⁵ With a similar mechanism of action, gentamicin exhibits its bactericidal function by binding to the 30S subunit, leading to functionally impaired proteins that ultimately result in cell death.⁶ The third antibiotic tested is cefotaxime sodium, which shows a high affinity for penicillin-binding proteins (PBPs), essential for bacterial wall biogenesis.⁷ When cefotaxime binds to the PBPs, it inactivates these proteins, hence weakening the bacterial cell wall and ultimately leading to bacterial lysis.⁸ The fourth antibiotic tested is chloramphenicol, and it binds to the 50S ribosomal subunit to interrupt the peptidyl transferase enzyme necessary for protein synthesis.¹⁰ By doing so, it halts the elongation of the growing polypeptide chain and prevents bacterial growth. Lastly, erythromycin also targets the 50S subunit, leading to slowed bacterial growth and consequently achieving a bacteriostatic effect.¹¹

Despite using antibiotics as medication to treat *S. aureus*, the emergence of drug-resistant strains like the Methicillin-resistant *Staphylococcus aureus* (MRSA) makes it difficult to treat the bacterial infection with single antimicrobials. MRSA is resistant to beta-lactam antibiotics, including the widely prescribed penicillins and cephalosporins, and it is directly associated with rising *S. aureus* death rates and lengthened hospital stays.¹² Due to overexposure to common antibiotics, bacterial strains can develop resistance mechanisms, including enzymatic inactivation of the antibiotic, efflux pumps, and decreased binding

affinity for the antibiotic, hence undermining the function of the antibiotics.¹³ Indeed, infections caused by antibiotic-resistant bacteria led to more than 33,000 deaths and 2.5 million additional days of hospitalization every year in the European Union, further resulting in an economic impact of nearly 900 million Euros.¹⁴ Standard antibiotics typically used to treat a wide range of bacterial pathogens are no longer as effective at treating many *S. aureus* strains; therefore, independent antimicrobial treatments may necessitate higher dosages to achieve substantial therapeutic effects. For this reason, using independent antibiotics to treat multidrug-resistant bacterial strains like *S. aureus* can lead to higher and quicker rates of bacterial resistance.

To combat this issue, combination therapy is a new approach that is being explored in treating *S. aureus* infections and other bacterial pathogens. It involves combining multiple antibiotics with varying mechanisms of action to achieve a synergetic effect, enhancing the therapeutic effectiveness compared to individual antibiotic treatment.¹⁵ The combination of antibiotics expands the treatment efficacy of the medication against bacteria, as multiple recombinant strains often cause pathogens. For example, beta-lactam antibiotics are commonly used against *Staphylococcus* and *Streptococcus* infections; however, if these antibiotics are administered in combination with Aminoglycosides, the efficacy of the treatment can expand to treating bacteria present in the intestines as well.¹⁶ Moreover, combination therapy often involves administering lower dosages of individual antibiotics for a significantly more effective pharmacological effect, leading to a wider range of possibilities for treatment. In many cases, these combinations decrease the time needed for recovery and treatment of bacterial infections like *S. aureus*. Further, research has shown that these multi-drug combinations can eradicate biofilm-forming pathogenic infections.¹⁷ Specifically, it has been shown that anprocide, a type of steroidal amine, is effective against the biofilms formed by *S. aureus* and *S. epidermidis*.¹⁸

To evaluate the effect of specific antibiotic combinations on *S. aureus* infections, we utilized the Newman strain of *S. aureus* and grew colonies exposed to varying concentrations of five different antibiotics. After analyzing the effects of the independent antibiotics on the Newman strain, we selected two of the most resistant antibiotics to use in combination treatment. We aimed to determine if the combination treatment would yield statistically significant results in inhibiting bacterial growth compared to individual antibiotic treatment. While previous studies have tackled the efficacy of independent antibiotics on strains of *S. aureus*, this paper aims to determine the effect of specific combinations on the Newman strain that have not been tested in detail in the past. Overall, this study aspires to contribute to the growing field of combination therapies for *S. aureus* and similar resistant strains.

■ Methods

1. Data and Methodology:

1.1. Recovery of *S. aureus* and Preparation of Antibiotics:

1.1.1. Recovery of Newman Bacteria:

In this study, we utilized the Newman strain of *S. aureus*, previously stored in a 4°C refrigerated environment. Initially, we added 2 mL of Tryptic Soy Broth (TSB) into a 14 mL polypropylene round-bottom tube. Using a sterile pipette, three clones of the Newman's strain were then swiped and carefully placed into the tube containing the 2 mL of TSB. The contents in the tube were swirled to ensure thorough mixing. We then incubated the tube in a 37°C warm culture room with shaking overnight to maximize aerobic bacterial growth and recovery.

1.1.2. Preparation of Antibiotic Solutions:

We used five antibiotics in this study, namely kanamycin, erythromycin, chloramphenicol, gentamicin, and cefotaxime. We measured out 100 mg each of solid erythromycin and cefotaxime sodium salt. We dissolved the 100 mg solid erythromycin in 2 mL of ethanol, while we dissolved 100 mg cefotaxime sodium salt in 2 mL of water, both prepared in separate 15 mL conical tubes over an open flame until the antibiotic was fully dissolved in the solution. The remaining three antibiotics were already in aqueous form and were used directly without any additional preparation. All five prepared antibiotic solutions were stored in a refrigerator at 4°C overnight before use. Before any dilutions, the kanamycin solution had an initial concentration of 40 mg/mL, while the other four antibiotic solutions (erythromycin, chloramphenicol, gentamicin, and cefotaxime) had initial concentrations of 50 mg/mL.

1.1.3. Bacterial Dilution and Counting:

Newman's bacterial density was measured using a Microplate Reader in terms of its Optical Density at 600 nanometers (OD600). We prepared a control cuvette by adding 1 mL of TSB over an open flame, while we added 1 mL of the recovered Newman bacteria into an experimental cuvette. The Microplate Reader was calibrated to the TSB control cuvette, which was set to zero. The starting overnight bacterial density showed expected day-to-day variation. Before the individual antibiotic experiment, the bacterial culture OD600 measured 2.094; before the combination antibiotic experiment, the OD600 measured 1.880.

An inoculum of approximately 10^8 CFU/mL was used, as this is a standard density for antibiotic susceptibility screening. This density provides a reliable OD600 reading and stays within the Microplate Reader's linear range once diluted in the 200 μ L wells. To achieve the standard concentration of 10^8 CFU/mL for the experiment, the bacterial sample needed to be diluted with Phosphate-Buffered Saline (PBS). We used the established conversion factor of $1 \text{ OD} = 5 \times 10^8 \text{ CFU}$. By rearranging terms, we found the dilution factors to be 10.47 and 9.40 for the individual antibiotic experiment and the combination antibiotic experiment, respectively.

We used OD600 as a unit of measurement in later experiments to determine the efficacy of either individual antibiotic

treatments or a combination of two antibiotics. Each experimental group contained three wells for the purpose of accuracy, all of which were read by the Microplate Reader to obtain OD600 values. All tables below detailing OD600 of experimental groups, such as Table 3, describe the average OD600 value among the three wells.

1.1.4. Dilution of Antibiotics.

To test the effect of different concentrations of antibiotics on the Newman strain of *S. aureus*, we diluted the antibiotics to concentrations of 1 µg/mL, 2.5 µg/mL, 5 µg/mL, 10 µg/mL, 25 µg/mL, 50 µg/mL, 100 µg/mL, and 500 µg/mL. The preparation of the dilutions was calculated, and the volumes of antibiotic and water are detailed in Tables 1 and 2, as shown below.

Table 1: Volume of kanamycin solution and water needed to create the respective diluted antibiotic concentrations.

Diluted Concentration (µg/mL)	Volume of Antibiotic (µL)	Volume of Water (µL)
1	5.00	995.00
2.5	12.50	987.50
5	2.50	975.00
10	5.00	95.00
25	12.50	87.50
50	25.00	75.00
100	50.00	50.00
500	2.50	N/A

Table 2: Volume of antibiotic (gentamicin, erythromycin, chloramphenicol, cefotaxime) solutions and water needed to create respective diluted antibiotic concentrations.

Diluted Concentration (µg/mL)	Volume of Antibiotic (µL)	Volume of Water (µL)
1	4.00	996.00
2.5	10.00	990.00
5	2.00	98.00
10	4.00	96.00
25	10.00	90.00
50	20.00	80.00
100	40.00	60.00
500	2.00	N/A

1.2. Broth Dilution Method:

1.2.1. Independent Antibiotic Treatment:

We added 200 µL of TSB over an open flame to 45 wells in a 96-well cell culture plate. In each well, 1 µL of the diluted Newman bacteria was added, followed by 1 µL of the diluted antibiotic concentration specific to each well. However, we did not need to dilute the antibiotics for the 500 µg/mL tests; instead, we added 2.5 µL of the initial kanamycin solution (40 mg/mL) and 2 µL for the other four antibiotics (50 mg/mL). We mixed the solutions in each well thoroughly with a pipette, and tips were changed after each well to prevent contamination. The culture plate was stored in a warm room at 37°C overnight for ample bacterial growth.

1.2.2. Combination Antibiotic Treatment:

Minimum Inhibitory Concentration (MIC) testing was performed to determine the antimicrobial susceptibility of *S. aureus* Newman when treated with the various individual antibiotic groups. MIC testing is performed when a standardized

amount of bacteria is exposed to serial dilutions of an antibiotic treatment group to establish the lowest concentration of the treatment that completely prevents visible growth of the bacteria. In our tests, we determined the MIC of each antibiotic treatment group to be the lowest treatment concentration without any bacterial growth.

After the MIC test with the independent antibiotic treatment was retrieved and evaluated with respective OD600 values, we determined the Newman strain to have the most resistance against kanamycin (at 25 µg/mL and 50 µg/mL), chloramphenicol (at 5 µg/mL and 10 µg/mL), and gentamicin (at 5 µg/mL and 10 µg/mL). In the procedure and results sections, we will use the initial letter of the antibiotic and the concentration value to represent the tested group. For instance, K25 refers to a treatment using 25 µg/mL concentrated kanamycin; similarly, K50/C5 refers to a treatment of 50 µg/mL concentrated kanamycin in combination with 5 µg/mL concentrated chloramphenicol.

Our combination therapy experiment consisted of 15 groups, with three wells allocated per group. This part of the experiment consisted of eight combination groups (K25/C5, K25/C10, K25/G5, K25/G10, K50/C5, K50/C10, K50/G5, K50/G10), six single antibiotic groups (K25, C5, G5, K50, C10, G10), and one negative control group with no antibiotic treatment. We repeated the dilution of the Newman bacterial sample, where the OD600 measurement showed a value of 1.880. We diluted the sample to 10^8 bacteria/mL, then added 1 µL of the diluted bacteria and 200 µL of TSB into each well. Finally, 1 µL of the respective antibiotics was added to the appropriate wells. The culture plate was stored at 37°C overnight for bacterial growth, and OD600 values were measured the following day to determine the efficacy of the treatments.

1.2.3. Experimental Controls:

Several control conditions were included in the experiments to ensure the results are valid, specifically through using sterility controls, positive controls for each antibiotic treatment group, and a sterile TSB control for the Microplate Reader calibration.

To ensure that the TSB growth medium was not contaminated with bacteria or other unwanted substances, we included three sterility control wells on each 96-well culture plate. Each sterility control well contained 200 µL TSB and 1 µL PBS with no bacteria present. After overnight incubation, any visible bacterial growth in these sterility control wells would suggest contamination of the growth medium and/or the pipettes.

For both the individual antibiotic and combination antibiotic treatments, each antibiotic row had positive control wells containing 200 µL TSB and 1 µL of the 108 CFU/mL diluted bacteria, with no antibiotic added to the wells. Positive control OD600 values were taken from the Microplate Reader to calculate percent reductions for the respective antibiotic treatments in varying concentrations.

We used cuvettes containing 200 µL of sterile TSB with no bacteria or antibiotic treatments to calibrate the Micro-

plate Reader. This sets the OD600 readings against a reference frame for more accurate readings.

■ Results and Discussion

1. Individual Antibiotic Treatment Results:

To identify independent antibiotic treatments resistant to some degree when treated to Newman *S. aureus*, we determined the Minimum Inhibitory Concentration (MIC) using the Broth Dilution Method, indicated by the lowest concentration of the antibiotic that leaves no visible trace of the bacteria following the incubation stage. The visibility of *S. aureus* growth was further confirmed with OD600 readings following the overnight incubation.

Results from the Microplate Reader determined that cefotaxime and erythromycin were both nearly immediately effective on Newman. With reference to Table 3, at just 1.0 µg/mL of aqueous cefotaxime administered, the OD600 reading dropped from 0.809 to 0.054. Similarly, 1.0 µg/mL of aqueous erythromycin administered led to an OD600 decrease from 0.688 to 0.067. Analysis of this data shows percentage decreases of 93.33% and 90.26% for cefotaxime and erythromycin after the initial dose, respectively. It can therefore be determined that Newman *S. aureus* exhibits little resistance to cefotaxime and erythromycin.

Comparatively, Table 3 indicates that the administration of chloramphenicol, kanamycin, and gentamicin was not effective at the initial doses. The MIC for chloramphenicol is between 5.0 µg/mL and 10.0 µg/mL: at a concentration of 5.0 µg/mL, we recorded an OD600 reading of 0.411, only a 37.92% decrease from the positive control; whereas, at a concentration of 10.0 µg/mL, we recorded an OD600 reading of 0.083, pointing to an 87.42% decrease from the positive control. Gentamicin exemplified similar trends, with an MIC between 5.0 µg/mL and 10.0 µg/mL: at a concentration of 5.0 µg/mL, we recorded an OD600 reading of 0.579, only a 24.21% decrease from positive control; conversely, at a concentration of 10 µg/mL, we found the OD600 reading to be 0.062, leading to a 91.88% decrease from the positive control group. Yet, out of all three antibiotics, Newman seemed to show the most resistance to kanamycin, as the MIC was indicated to be between 25.0 µg/mL and 50 µg/mL. The OD600 readings and percentage changes from the positive controls are calculated to be 0.401 and -39.33% for 25.0 µg/mL kanamycin, and 0.073 and -92.70% for 50 µg/mL kanamycin.

Table 3: Average OD600 results from individual antibiotic MIC treatment. Results found that individual administration of chloramphenicol, kanamycin, and gentamicin was not effective at initial doses; therefore, Newman showed a degree of resistance to these three antibiotics.

Antibiotic Type	Administered Concentration (µg/mL)								
	Control	1.0	2.5	5.0	10.0	25.0	50.0	100.0	500.0
Cefotaxime	0.809	0.054	0.055	0.053	0.051	0.051	0.050	0.050	0.054
Erythromycin	0.688	0.067	0.071	0.070	0.065	0.059	0.057	0.056	0.059
Chloramphenicol	0.662	0.517	0.466	0.411	0.083	0.050	0.051	0.051	0.053
Kanamycin	0.661	0.592	0.555	0.525	0.431	0.401	0.073	0.060	0.053
Gentamicin	0.764	0.605	0.578	0.579	0.062	0.057	0.054	0.055	0.053

Interpretation of the percent reduction values requires consideration of the variation in the positive control groups. Even within the same experiment, positive control OD600 values varied across different antibiotics (ranging from 0.661 to 0.809), which likely reflects minuscule differences in the inoculum while setting up separate rows of the 96-well plate. However, because each antibiotic was tested with its own positive control group, we calculated the OD600 percent reductions using the control group specific to the antibiotic rather than using a single baseline control. Between the individual antibiotic and combination antibiotic experiments, positive controls also differed slightly, likely due to normal day-to-day variation. Despite this, all wells—meaning the positive control groups, individual antibiotic treatments, and combination antibiotic treatments—all received the same starting bacterial density, so comparisons with the data remain valid.

2. Combination Antibiotic Treatment Results:

To determine the effect of combination antibiotic therapy on resistant Newman *S. aureus*, we combined kanamycin (25 µg/mL and 50 µg/mL), gentamicin (5 µg/mL and 10 µg/mL), and chloramphenicol (5 µg/mL and 10 µg/mL) to create eight experimental groups. OD600 readings for the combination antibiotic treatment using the Microplate Reader were collected and graphed using GraphPad Prism. The OD600 readings for all eight experimental groups were graphed alongside the positive control and their respective independent antibiotics. For statistical analysis, we ran unpaired two-tailed t-tests with 95% confidence intervals for each comparison (combination vs. two individual antibiotics, combination vs. control). The error bars in the graphs show the standard deviation (SD) of the three repeated wells per antibiotic treatment. The statistical significance in the analysis was defined with p-value thresholds (*p < 0.05, **p < 0.01, ***p < 0.001). Only Figures 1 and 2 are provided in this research paper to demonstrate the most remarkable combination treatment results.

Following statistical analysis, we determined the treatment administering gentamicin at 10 µg/mL combined with kanamycin at 50 µg/mL (G10/K50) to be most effective. Figure 1 indicates the combination group showed no statistical significance (p > 0.05) compared to the individual antibiotic treatments of gentamicin at 10 µg/mL and kanamycin at 50 µg/mL. However, the combination group compared to the positive control showed three asterisks (***p < 0.001), while both the individual antibiotic treatments showed only one asterisk (p < 0.05) compared to the positive control group. The statistical significance of the G10/K50 combination group can be supported with the data from the OD600 readings. The means of three tested wells for each antibiotic treatment are as follows: 0.3220 for gentamicin (10 µg/mL) only, 0.2907 for kanamycin (50 µg/mL) only, 0.05700 for combination treatment, and 0.6460 for positive control. Compared to the positive control group, gentamicin at 10 µg/mL showed a 50.15% decrease, kanamycin at 50 µg/mL showed a similar 55.00% decrease, while the combination treatment showed a drastic 91.18% decrease. Compared to the independent treatments, the change in percentage comparing the combination group to the indepen-

dent antibiotic treatments was calculated to be -82.30% and -80.39% for gentamicin (10 µg/mL) and kanamycin (50 µg/mL), respectively.

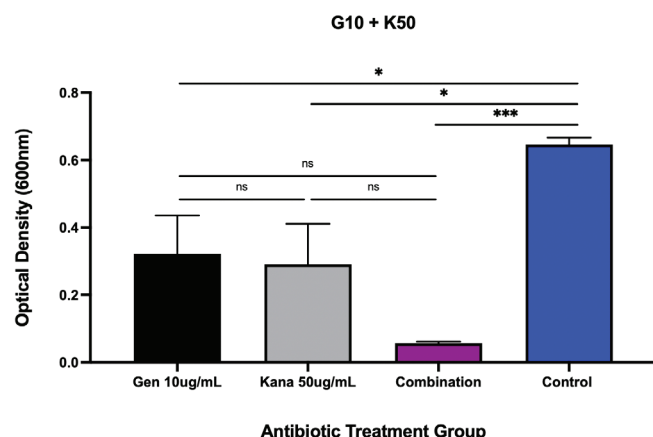


Figure 1: Results from combination therapy (gentamicin 10 µg/mL and kanamycin 50 µg/mL). Significance levels showed that this combination is the most effective treatment group in all eight combination groups.

From the OD600 readings and the analysis from the graph, we determined the second most effective combination treatment to be gentamicin (10 µg/mL) working with kanamycin (25 µg/mL). Figure 2 indicates that the combination group showed no significance ($p > 0.05$) compared to the individual 10 µg/mL gentamicin treatment, while it showed a significance of one asterisk ($p < 0.05$) compared to the individual 25 µg/mL kanamycin treatment. In comparison to the positive control group with no antibiotic treatment, the combination treatment showed a significance value of three asterisks ($p < 0.001$), while the individual kanamycin and gentamicin treatments showed no significance ($p > 0.05$) and one asterisk ($p < 0.05$), respectively. Based on the OD600 readings collected, the percent differences between the combination groups, the individual antibiotic treatment groups, and the positive control support the relationship from the graph. The OD600 mean values are as follows for the G10/K25 combination group: 0.3200 for gentamicin (10 µg/mL) only, 0.4800 for kanamycin (25 µg/mL) only, 0.08400 for combination treatment, and 0.6460 for the positive control group. Compared to the positive control group, gentamicin at 10 µg/mL showed a 50.15% decrease, kanamycin at 25 µg/mL showed a lower 25.70% decrease, while the combination treatment showed a notable 87.00% decrease, albeit being less than the 91.18% decrease from the positive control recorded from the G10/K50 combination group. Compared to the independent treatments, the change in percentage comparing the combination group to the independent antibiotic treatments was calculated to be -73.91% and -82.50% for gentamicin (10 µg/mL) and kanamycin (25 µg/mL), respectively. As aforementioned, while the results from the G10/K25 group are significant, the G10/K50 combination treatment shows statistically more significant results.

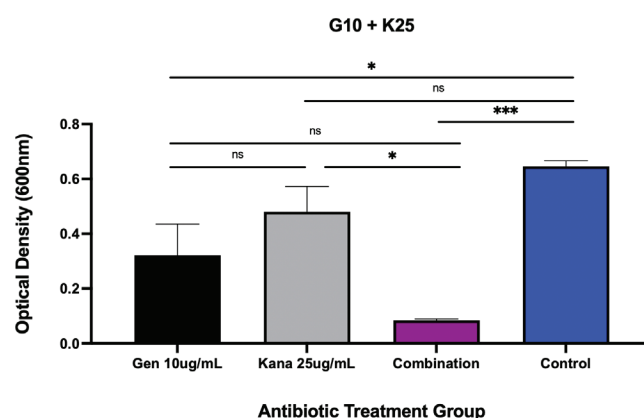


Figure 2: Results from combination therapy (gentamicin 10 µg/mL and kanamycin 25 µg/mL). Significance levels showed that this combination is the second most effective treatment group in all eight combination groups.

3. Discussion:

Our study demonstrated that, among eight combination therapies tested against *S. aureus* Newman, the combination of gentamicin (10 µg/mL) and kanamycin (50 µg/mL) exhibited the most significant antibacterial result. This was determined by analyzing the highest percentage of growth inhibition of the combination group compared to the positive control group and the two individual antibiotic groups.

The significant effect demonstrated by the G10/K50 combination therapy group could be attributed to a number of reasons. Both gentamicin and kanamycin are aminoglycoside antibiotics; their primary mechanism of action is to bind to the bacterial 30S ribosomal subunit, causing a misreading of the mRNA and further inhibition of protein translocation. Ultimately, this leads to bacterial cell death. Aminoglycoside antibiotics are also concentration-dependent, meaning higher peak concentrations of the doses lead to greater efficacy. It is therefore made clear that combining a high MIC of gentamicin and kanamycin, two drugs from the same class, led to an enhanced effect, meaning the results from combination therapy are significantly stronger compared to the results from individual therapy. We did not calculate a Fractional Inhibitory Concentration Index (FICI), which is the standard for defining formal synergy. Thus, we describe the effect we observed as “enhanced” rather than true synergy.

Since aminoglycosides are positively charged and cannot easily penetrate through the negatively charged cell membrane of *S. aureus*, their uptake into the cell is through a proton motive force generated by the bacterial electron transport chain. By creating a strong electrochemical gradient, the electron transport chain transports the aminoglycoside molecules into the cell. This process can help explain why the combination of gentamicin (10 µg/mL) with kanamycin (50 µg/mL) worked best against the formerly resistant *S. aureus* Newman. Understanding the relationship and mechanism of action for aminoglycosides, we propose that the combination of the two drugs could have created a positive feedback loop to enhance the overall antibacterial effect. The first drug, such as gentamicin at 10 µg/mL, could have initially entered a few bacterial cells. Once inside the cell, the gentamicin will begin to take

effect and create mistranslated proteins, some of which will insert themselves into the cell membrane. Having faulty proteins in the cell membrane could lead to lower permeability, making it much easier for the kanamycin at 50 µg/mL to enter the bacterial cells and cause even more rapid protein disruptions, therefore leading to more notable cell death compared to one individual antibiotic working alone. The dose of the antibiotics also plays an integral role; the concentration of the gentamicin at 10 µg/mL was enough to initiate the membrane-damaging process without being too high to work entirely well on its own, and likewise, the concentration of kanamycin at 50 µg/mL was enough to work well with the gentamicin to speed up the process of protein disruptions and bacterial cell death.

It is worthwhile to consider whether these antibiotic concentrations are clinically realistic and relevant. Gentamicin at 10 µg/mL is clinically possible, but it rests at the upper end of peak concentrations in humans, which typically ranges from 5 µg/mL to 12 µg/mL, depending on the dosing. However, kanamycin at 50 µg/mL exceeds its typical peak concentrations, ranging from 20 µg/mL to 30 µg/mL. The two antibiotics, gentamicin and kanamycin, are both aminoglycosides, a class of bactericidal antibiotics known to cause drug-induced nephrotoxicity.⁹ Put differently, aminoglycosides risk causing kidney damage and hearing loss at high doses. Hence, while the G10/K50 combination worked well *in vitro*, clinical use would need lower doses of the combination therapy to prevent drug-induced nephrotoxicity.

While the gentamicin at 10 µg/mL combined with kanamycin at 50 µg/mL showed the most significant antibacterial effects, all other experimental combination therapy groups also led to decreases in OD600 readings compared to individual therapies, albeit less notable. For the gentamicin and kanamycin pairs, data collected from the experiments suggest that the doses of either antibiotic below the critical threshold would lead to a less desired outcome. From the G10/K50 therapy group, in contrast to the G10/K25 group, it becomes clearer that most of the differences in the outcome can be highly attributed to the lower concentration of kanamycin. Most notably, all combinations involving chloramphenicol were significantly less effective. We propose that this could be due to a conflict in the mechanism of action, where the bacteriostatic action of chloramphenicol inhibits the metabolic activity needed for the uptake of aminoglycosides like kanamycin into the bacterial cell. It can therefore be made clearer why the G10/K50 combination seemingly worked most successfully among the eight combination therapy groups.

4. Limitations of the Study:

While this study provides evidence of an enhanced effect occurring between gentamicin and kanamycin against *S. aureus* Newman, several limitations should be considered when interpreting the results from the experiments.

Firstly, the experiments from this study were conducted utilizing a single strain of *S. aureus*, namely the Newman strain. The strain we utilized was extracted from patients with *S. aureus* and recovered before use in this study. The experiments with antibiotic treatment were administered *in vitro*, meaning

the effects of the combination therapy may not translate completely when the same treatments are applied *in vivo*. There are many factors, including host immune responses, that are important in affecting antibiotic efficacy yet difficult to replicate outside of the patient's body. Therefore, it is made more difficult to determine whether the results from this lab still stand valid under *in vivo* conditions.

Moreover, while the Newman strain is widely used in biomedical research to test for *S. aureus* trends, testing on only the Newman strain cannot be representative of all the different isolates and strains under the generalized term *Staphylococcus aureus*. Thus, it can be difficult to ascertain whether these results are merely due to variance in this one strain or if they remain true under different circumstances. In addition, each antibiotic treatment group was tested in only one independent experiment with three replicate wells. We did not repeat the entire experiment on different days, which limits our ability to assess whether the results would remain valid on different days with new bacterial cultures. Assuming that our proposed mechanism is true for G10/K50, the entry into the bacterial cell of the first aminoglycoside (either gentamicin or kanamycin) is a process that can be affected by factors including the presence of specific protein transporters, variance in the electron transport chain, and membrane permeability. Without testing the combination therapy across a variety of strains, including MRSA and strains with known aminoglycoside resistance, the results from this study are limited.

Lastly, our proposed mechanism of action for the G10/K50 combination group is merely an inference based on current research and the observed results. We predicted that the success in the G10/K50 combination could be due to gentamicin enhancing the effect and uptake of kanamycin through the initial protein disruption in the membrane; however, our study did not directly test this hypothesis. Without confirmation of this mechanism, it is difficult to conclusively explain the results from this study, regardless of how plausible this explanation may be.

5. Future Directions:

Future studies can build on these findings in a number of ways. To begin, checkerboard assays can be run to determine the Fractional Inhibitory Concentration Index (FICI) for the G10/K50 and G10/K25 antibiotic combinations, which would confirm whether the enhanced bactericidal effect can be deemed synergistic using its formal definition. Additionally, the G10/K50 combination should be tested on other *S. aureus* strains, including MRSA, to see if the strong bactericidal effect works beyond the Newman strain of *S. aureus*. Membrane permeability assays, such as propidium iodide staining, can also directly test our proposed mechanism that one aminoglycoside influences and amplifies the absorption of the other aminoglycoside in the combination treatment. These future experiments would confirm some of the mechanisms we proposed and minimize any uncertainties.

■ Conclusion

In conclusion, this study identified the combination of gentamicin at 10 µg/mL and kanamycin at 50 µg/mL to be the most effective treatment against the Newman strain of *Staphylococcus aureus* in the laboratory *in vitro*, showing an enhanced effect that ultimately leads to notable bacterial cell death. In the initial stage, *S. aureus* Newman showed some resistance to gentamicin, chloramphenicol, and kanamycin until the dose reached a high enough threshold; contrarily, Newman showed little to no resistance to cefotaxime and erythromycin. The combination therapy is shown to have a stronger effect in comparison to individual antibiotic treatments, and even more notable when compared to the *S. aureus* group exposed to no antibiotic treatment. We propose that the combination therapy of gentamicin at 10 µg/mL and kanamycin at 50 µg/mL is most successful due to a mechanism that allows for increased antibiotic uptake in the bacterial cell, leading to more rapid malfunctions inside the cell over a shorter period of time. Despite clinical relevance being limited by this study from its *in vitro* nature and use of only one strain of *S. aureus*, the findings from this study still highlight notable results about the efficacy of certain combination therapy groups over individual treatment for the five original individual antibiotics.

■ Acknowledgments

I would like to give my sincere appreciation to PhD student Ms. Zhao for her relentless guidance, feedback, and inspiration throughout this entire research process. I would also like to give my immense thanks to Prof. Huang for providing the lab materials and for his genuine support in this project, and to my family for their utmost encouragement throughout this entire process.

■ References

- van Hal, S. J.; Jensen, S. O.; Vaska, V. L.; Espedido, B. A.; Paterson, D. L.; Gosbell, I. B. Predictors of Mortality in *Staphylococcus Aureus* Bacteremia. *Clinical Microbiology Reviews* **2012**, *25* (2), 362–386. <https://doi.org/10.1128/cmr.05022-11>.
- Etter, D.; Corti, S.; Spirig, S.; Cernela, N.; Stephan, R.; Johler, S. *Staphylococcus Aureus* Population Structure and Genomic Profiles in Asymptomatic Carriers in Switzerland. *Frontiers in Microbiology* **2020**, *11*. <https://doi.org/10.3389/fmicb.2020.01289>.
- Taylor, T. A.; Unakal, C. G. *Staphylococcus aureus* infection. *National Library of Medicine*. <https://www.ncbi.nlm.nih.gov/books/NBK441868/>.
- Centers for Disease Control and Prevention. *Staph infections can kill*. Centers for Disease Control and Prevention. <https://www.cdc.gov/vitalsigns/staph/index.html>.
- Bruno, A. Identifying Patterns on the Development of Kanamycin Resistance in *Staphylococcus epidermidis* with Four-Welled Plates. <https://doi.org/10.13140/RG.2.2.16783.20649>.
- Chaves, B. J.; Tadi, P. *Gentamicin*. PubMed. <https://www.ncbi.nlm.nih.gov/books/NBK557550/>.
- Drugbank. *Cefotaxime*. [go.drugbank.com](https://go.drugbank.com/drugs/DB00493). <https://go.drugbank.com/drugs/DB00493>.
- Macheboeuf, P.; Fischer, D. S.; Brown, T.; Zervosen, A.; Luxen, A.; Joris, B.; Dessen, A.; Schofield, C. J. Structural and Mechanistic Basis of Penicillin-Binding Protein Inhibition by Lactvicins. *Nature Chemical Biology* **2007**, *3* (9), 565–569. <https://doi.org/10.1038/nchembio.2007.21>.

- Mingeot-Leclercq, M. P.; Tulkens, P. M. Aminoglycosides: Nephrotoxicity. *Antimicrobial Agents and Chemotherapy* **1999**, *43* (5), 1003–1012. <https://doi.org/10.1128/AAC.43.5.1003>.
- What is the mechanism of Chloramphenicol? Patsnap.com. <https://synapse.patsnap.com/article/what-is-the-mechanism-of-chloramphenicol>.
- What is the mechanism of Erythromycin? Patsnap.com. <https://synapse.patsnap.com/article/what-is-the-mechanism-of-erythromycin>.
- Methicillin-Resistant *Staphylococcus Aureus* (MRSA) – How Can People Become Infected? 2025. <https://www.bfr.bund.de/cm/349/questions-and-answers-about-methicillin-resistant-staphylococcus-aureus.pdf> (accessed 2026-01-05).
- Pantosti, A.; Sanchini, A.; Monaco, M. Mechanisms of Antibiotic Resistance In *Staphylococcus Aureus*. *Future Microbiology* **2007**, *2* (3), 323–334. <https://doi.org/10.2217/17460913.2.3.323>.
- Ratia, C.; Soengas, R. G.; Soto, S. M. Gold-Derived Molecules as New Antimicrobial Agents. *Frontiers in Microbiology* **2022**, *13*. <https://doi.org/10.3389/fmicb.2022.846959>.
- Kamble, E.; Sanghvi, P.; Paredesi, K. Synergistic Effect of Antibiotic Combinations on *Staphylococcus Aureus* Biofilms and Their Persister Cell Populations. *Biofilm* **2022**, *4*, 100068. <https://doi.org/10.1016/j.bioflm.2022.100068>.
- When is it necessary to combine antibiotics? Vinmec International Hospital. <https://www.vinmec.com/eng/blog/when-to-combine-antibiotics-en>.
- Bognár, B.; Spohn, R.; Lázár, V. Drug Combinations Targeting Antibiotic Resistance. *npj Antimicrobials and Resistance* **2024**, *2* (1). <https://doi.org/10.1038/s44259-024-00047-2>.
- Pettit, R. K.; Weber, C. A.; Lawrence, S. B.; Pettit, G. R.; Kean, M. J.; Cage, G. D. *In Vivo* Activity of Anprocide Alone, and *In Vitro* Activity in Combination with Conventional Antibiotics against *Staphylococcus Aureus* and *Staphylococcus Epidermidis* Biofilms. *Journal of Medical Microbiology* **2009**, *58* (9), 1203–1206. <https://doi.org/10.1099/jmm.0.008268-0>.

■ Author

Olivia Liu is a senior at Hong Kong International School. Passionate about the intersection of biomedical research and healthcare to improve public health for all, she hopes to continue conducting meaningful research and contributing to medicine, with plans to pursue a pre-med track in her future collegiate studies..

A Systematic Review of Takotsubo Cardiomyopathy in Oncologic Cohorts

Pranav Maddineni

Farmington High School, 26 Baldwin Dr, Farmington, Connecticut, 06032, USA; maddinenipranav@gmail.com

Mentor: Prahlad Parajuli

ABSTRACT: Patients with cancer experience significant emotional, physical, and inflammatory stress, which can lead to Takotsubo cardiomyopathy (TCM), a reversible but acute type of ventricular failure. Understanding how TCM manifests in oncology is essential, as it frequently resembles acute coronary syndrome and overlaps with chemotherapy-related toxicity. In this systematic review of case reports, series, and registry data, we found that both malignancy itself and anti-neoplastic therapies are credible precipitants of TCM. Across more than 40 published cases, patients typically presented with chest pain, ST-segment changes, modest troponin elevation, and unobstructed coronaries. Imaging revealed transient apical ballooning and recovery of ejection fraction within weeks of drug withdrawal. Severe complications such as ventricular arrhythmias, thrombus formation, and embolic stroke were occasionally reported, and rechallenge with fluoropyrimidines carried a risk of recurrence. Overall, this review synthesized case-level evidence to clarify the presentation and mechanisms of TCM in oncology patients, with three main findings: (1) diagnostic features and clinical outcomes, (2) anti-neoplastic drugs most consistently associated with TCM, and (3) management strategies, including risks noted with rechallenge. These findings suggest that oncologists and cardiologists should maintain a low threshold for suspecting TCM in cancer patients presenting with ACS-like features, particularly during early cycles of fluoropyrimidines, HER2-targeted agents, and immune checkpoint inhibitors. Future research should prioritize prospective registries and mechanistic studies to guide evidence-based cardioprotective strategies in this vulnerable population.

KEYWORDS: Translational Medical Sciences, Disease Detection and Diagnosis, Cardio-oncology, Takotsubo Cardiomyopathy, Anti-neoplastic Agents.

■ Introduction

Cardio-oncology faces a persistent challenge in distinguishing chemotherapy-induced cardiotoxicity from other cardiac events in patients undergoing systemic anti-neoplastic therapies. Many of these patients already carry cardiovascular risk factors or preexisting disease, while simultaneously experiencing substantial physical and emotional stress during diagnosis and treatment.¹ Misattributing cardiac injury can have serious consequences: failing to recognize drug-related toxicity risks progressive, potentially irreversible damage, whereas wrongly discontinuing an effective anti-cancer agent for a non-drug cardiac event may compromise oncologic outcomes and survival.^{1,2} This clinical uncertainty highlights a pressing research gap in understanding stress-related cardiomyopathies within oncology.

Takotsubo cardiomyopathy (TCM) exemplifies this gap. It is a reversible, stress-related left-ventricular dysfunction that mimics *Acute Coronary Syndrome* (ACS) with chest pain, ECG changes, and modest troponin elevation but no obstructive coronary disease on angiography.^{1,2} The prototypical imaging pattern is transient apical ballooning with basal hyperkinesis.³ Although uncommon in the general population, registry and pooled analyses indicate that cancer is disproportionately represented among TCM cohorts, with prevalence estimates of 10–17% in modern series.⁴ Importantly, oncology-associated TCM appears more often linked to physical stressors and

carries worse short-term outcomes than emotionally triggered cases.^{2,4}

Case reports and small series further support this association, implicating several anti-neoplastic agents as temporal triggers of TCM. Fluoropyrimidines such as 5-fluorouracil and capecitabine are the most consistently reported, but targeted therapies, including trastuzumab and bevacizumab, have also been reported.^{1,5,6} These observations suggest that both cancer itself and its treatments act as precipitating factors.

Despite many reports, important uncertainties remain. Many reports lack complete diagnostic workups (for example, cardiac MRI or biopsy to exclude myocarditis), so attributing a single chemotherapeutic drug as the cause is therefore uncertain.^{2,6} Researchers caution that the psychological stresses of cancer care itself could induce TCM-like syndromes, so we must be careful before labeling an agent as definitively causative.² Taken together, the literature supports an association between cancer care and TCM, but also reveals inconsistent diagnostic completeness and variations in the reporting of timing, severity, and outcomes.^{1,2}

To address these gaps, this study compiles over 40 published case reports into a structured spreadsheet, creating an international case-level registry analysis. The review offers a unified viewpoint on Takotsubo cardiomyopathy in oncology by methodically extracting clinical presentation, diagnostic results, treatment approaches, and outcomes. The purpose of this synthesis is to shed light on diagnostic trends, identify the

drug classes most frequently involved, and guide useful management choices in cardio-oncology.

■ Methods

This study examined the prevalence and features of Takotsubo cardiomyopathy (TCM) in cancer patients, with an emphasis on cases associated with anti-neoplastic treatments. The objective was to compile disparate case-level material into an organized study that could shed light on management tactics, implicated agents, and diagnostic characteristics. This work was conducted as a literature review rather than an in-field clinical study. The dataset was built from published case reports, case series, registry analysis, and previous reviews. Keyword combinations like “Takotsubo Cardiomyopathy,” “TCM,” “TTS,” “Stress-Cardiomyopathy,” “apical ballooning,” “cancer,” “chemotherapy,” “antineoplastic,” and agent-specific terms like “5-fluorouracil,” “capecitabine,” “trastuzumab,” “bevacizumab,” and “immune checkpoint inhibitor” were used to collect data from PubMed/MEDLINE and ClinicalTrials.gov. Titles and abstracts were screened, full texts reviewed, and reference lists hand-searched to capture additional cases. Over 40 published case reports and series were extracted into a standardized spreadsheet documenting patient demographics, cancer type, drug exposure and timing, diagnostic findings, management, and outcomes, allowing comparison across cases and identification of recurring patterns. All information was derived from peer-reviewed sources, and no laboratory or clinical experimentation was performed; accordingly, no ethical approval or human or animal subject considerations were required.

Overview of Takotsubo Cardiomyopathy :

Takotsubo Cardiomyopathy (TCM), first characterized in Japan in the early 1990s, is a transient syndrome of left ventricular systolic dysfunction that commonly presents like an *Acute Coronary Syndrome* (ACS) with chest pain, ischemic electrocardiographic (ECG) changes, and modest troponin rise, yet without *Obstructive Coronary Disease* on angiography.^{1,3} The term “takotsubo” refers to the Japanese octopus trap-like shape formed by akinetic apical segments on ventriculography, while the basal segments remain hyperkinetic.³ Clinically, patients may present with chest pain, dyspnea, syncope, or hemodynamic compromise; complications can include heart failure, arrhythmia, left ventricular thrombus, embolic events, cardiogenic shock, and, very rarely, death.^{1,5}

Epidemiology and Clinical Patterns :

TCM is uncommon in the general population but is well documented across large registries and cohort studies. Modern registries show that over 90% of the typical patient is an older woman, predominantly in the postmenopausal age group, and many cases follow an identifiable emotional or physical stressor.^{5,7,8} As said before, TCM accounts for approximately 1–3% of patients presenting with a clinical picture suggestive of ACS.^{1,7} Emotional or physical stressors are identified in about two-thirds of cases, though up to one-third have no apparent precipitant.⁸ The InterTAK registry highlighted that

neurologic disease and psychiatric disorders are more prevalent in patients with TCM compared with other Acute Coronary Syndromes, suggesting a role for neurohumoral and stress-mediated pathways.⁸

Pathophysiology:

The pathophysiology of TCM is multifactorial, with sympathetic nervous system activation at its core. Early descriptions (1990s) by Dote and colleagues reported patients with multivessel coronary vasospasm producing reversible myocardial dysfunction, suggesting that epicardial spasm may explain a subset of cases.³ However, the pivotal study by Wittstein *et al.* demonstrated markedly elevated plasma catecholamine levels in patients with TCM, several-fold higher than in patients with ST-elevation MI, supporting a central catecholamine-mediated injury.⁹ At the myocardial level, excess catecholamines produce intracellular calcium overload, cAMP-mediated metabolic stress, and maladaptive β -adrenergic signaling, including a proposed β_2 -adrenergic switch to Gi coupling, changes that can provoke reversible myocyte stunning and regional wall-motion abnormalities characteristic of TCM.^{1,8,9} Coronary microvascular dysfunction has been shown by multimodality imaging and physiological studies, and likely causes transient tissue hypoperfusion even when epicardial arteries are angiographically normal.^{8,10} Early reports of simultaneous multivessel epicardial spasm indicate that coronary vasospasm explains a minority of cases in which spasm is demonstrable or provoked. The strong predominance of the postmenopausal age group of women and older patients points to sex- and age-related modifiers of susceptibility, such as loss of estrogenic cardioprotection and altered adrenergic or endothelial responsiveness.^{4,8} Genetic and molecular predispositions have been proposed but remain incompletely defined and definitely need further investigation.^{4,8} Figure 1 illustrates the prevailing model, in which an acute trigger induces a catecholamine surge that injures a susceptible myocardium through direct myocyte effects, microvascular dysfunction, and occasionally, epicardial spasm, producing the transient regional systolic failure seen in TCM.^{1,3,8,9,10}

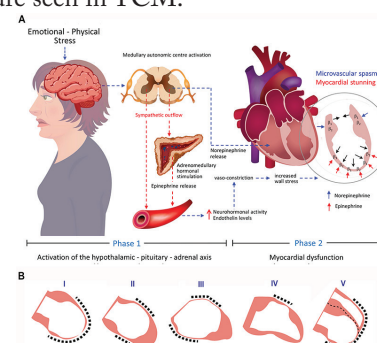


Figure 1: Diagram illustrating the pathophysiology of *Takotsubo Cardiomyopathy*. A catecholamine surge triggered by acute emotion or physical stimuli acts on β_1 - and β_2 -adrenergic receptors, producing direct myocardial injury, intracellular calcium overload, cAMP-mediated stress, and regional wall-motion abnormalities. These processes produce the characteristic left-ventricular contraction pattern, with apical akinesis or dyskinesis and basal hyperkinesis, which defines the typical presentation of TCM. Image reproduced from: Díaz-Navarro, R. Takotsubo Syndrome: The Broken-Heart Syndrome. *British Journal of Cardiology* 2021. <https://doi.org/10.5837/bjc.2021.011>.

Diagnostic Considerations:

Clinically, TCM must be distinguished from *Acute Myocardial Infarction*. Angiography remains essential to exclude obstructive coronary disease.⁸ Multimodality imaging, including ECG and cardiac magnetic resonance, refines the diagnosis by characterizing ventricular morphology, ruling out myocarditis, and detecting complications such as thrombus. The Mayo Clinic criteria remain widely used for case definition and research purposes.¹¹ Common diagnostic findings are summarized below in Figures 2 and 3.

Mayo Clinic Diagnostic Criteria for Takotsubo Cardiomyopathy:¹¹

- 1. Transient hypokinesis, akinesis, or dyskinesis of left ventricular mid-segments with/without apical involvement, regional wall motion abnormality extending beyond a single vascular territory, and a stressful trigger are often, but not always, present;
- 2. Absence of angiographic evidence of obstructive coronary disease or plaque rupture;
- 3. New ECG abnormalities consisting of ST elevation and/or T wave inversion with modest elevation in cardiac troponins;
- 4. Absence of pheochromocytoma and myocarditis

Electrocardiographic changes in TCM typically follow a dynamic pattern, evolving from initial ST-segment elevation to T-wave inversion and QT prolongation over days to weeks. Figure 2 illustrates these characteristic ECG alterations and their differential value relative to true myocardial infarction.

ST elevation	On presenting ECG • 42-46% in precordial leads • Concomitant inferior or lateral leads • Uncommon if isolated in inferior or lateral leads	
ST depression	On presenting ECG • Potential findings but less frequent (8%)	
Branch Block	On presenting ECG in 5% • Left • Right	
T waves inversion	Evolution of ECG in 41% • Diffuse and symmetric • Slow resolution and often partial resolution	
Pathologic Q waves	• 6-31% • Transient Q waves are unusual	
QT segment	On presenting ECG • Prolonged up to 500ms • Normalises in few days	
PR segment	On presenting ECG • Prolonged	

Figure 2: Electrocardiographic (ECG) alterations are commonly observed in Takotsubo cardiomyopathy (TCM). These changes often mimic acute coronary syndromes, with ST-segment elevation in precordial leads being the most frequent finding. At the same time, ST depression, bundle branch block, and transient Q waves occur less often. T-wave inversion and QT prolongation are characteristic of the recovery phase, evolving over days to weeks, and may predispose patients to arrhythmias. To differentiate TCM from myocardial infarction and provide appropriate treatment, it is crucial to recognize this dynamic pattern of ECG anomalies. Image reproduced from: Pillitteri, M.; Brogi, E.; Piagnani, C.; Bozzetti, G.; Forfori, F. Perioperative Management of Takotsubo Cardiomyopathy: An Overview. *Journal of Anesthesia, Analgesia and Critical Care* 2024, 4 (1). <https://doi.org/10.1186/s44158-024-00178-y>.

Imaging is central to confirming the diagnosis of TCM, with coronary angiography used to exclude obstructive coronary disease and left ventriculography documenting the hallmark apical ballooning pattern. Figure 3 presents representative angiographic and ventriculographic findings from a typical TCM case, with arrows indicating the absence of obstructive lesions and panels demonstrating systolic apical ballooning with basal hyperkinesis.

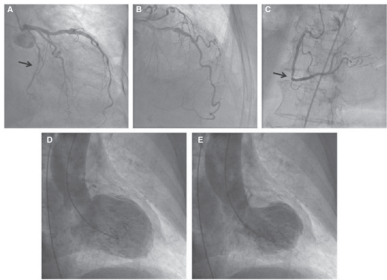


Figure 3: Typical TCM imaging. Coronary angiography in Panels A–C reveals no obstructive coronary artery disease (angiographic views are indicated by arrows). Panels D–E: left Ventriculography showing relative basal hyperkinesis and systolic apical ballooning. Image reproduced from: Citro, R.; Okura, H.; Ghadri, J. R.; Izumi, C.; Meimoun, P.; Izumo, M.; Dawson, D.; Kaji, S.; Eitel, I.; Kagiya, N.; Kobayashi, Y.; Templin, C.; Delgado, V.; Nakatani, S.; Popescu, B. A.; Bertrand, P.; Donal, E.; Dweck, M.; Galderisi, M.; Haugaa, K. H. Multimodality Imaging in Takotsubo Syndrome: A Joint Consensus Document of the European Association of Cardiovascular Imaging (EACVI) and the Japanese Society of Echocardiography (JSE). *European Heart Journal - Cardiovascular Imaging* 2020, 21 (11), 1184–1207. <https://doi.org/10.1093/ehjci/jeaa149>.

TTS And Cancer :

Cancer, a complex and long-term condition, can heighten the risk of Takotsubo Cardiomyopathy (TCM) due to the intense emotional and physical stress it causes. Emotional strain from diagnosis, chronic inflammation seen in many malignancies, and the physical toll of the disease or surgery are all recognized contributors to TCM.¹ Some studies also suggest that paraneoplastic mediators may interfere with cardiac adrenoceptors, potentially leading to contractile dysfunction.¹⁴ Cancer-related pain crises can also contribute to TCM development. The prevalence of malignancy in patients diagnosed with TTS is substantial, ranging from 4% to 29% across various studies, with some registries reporting rates as high as 15.8%. This indicates a notable overrepresentation of cancer patients within TCM cohorts when compared to the general population.¹⁴ These overlapping factors such as emotional distress, inflammation, treatment-related stressors, and paraneoplastic signaling are visually illustrated in Figure 4 (Figure 4).

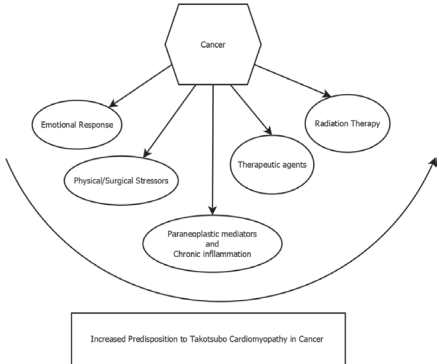


Figure 4: Through several interrelated processes, cancer may increase susceptibility to Takotsubo cardiomyopathy. While the disease itself and associated surgical treatments cause severe physical strain, the diagnosis and progression of cancer can cause substantial mental distress. Cardiovascular function can also be changed by the paraneoplastic mediators linked to cancer and the chronic inflammatory environment. Finally, various cancer treatments, including chemotherapy, targeted therapies, and radiation, may act as direct physiological triggers for the onset of TTS. Image reproduced from: Desai, A.; Noor, A.; Joshi, S.; Kim, A. S. Takotsubo Cardiomyopathy in Cancer Patients. *Cardio-Oncology* 2019, 5 (1). <https://doi.org/10.1186/s40959-019-0042-9>.

Anti-Cancer Therapy-Induced Cardiotoxicity:

The cardiotoxicity of various chemotherapeutic agents is a well-established concern in cardio-oncology.¹⁶ Cancer therapies can elicit a broad spectrum of adverse cardiac effects, including left ventricular dysfunction, myocardial ischemia, arrhythmias, hypertension, accelerated atherosclerosis, venous thromboembolism, myocarditis, and QT prolongation.¹⁷⁻²⁰ These complications restrict the use of certain chemotherapy drugs and represent a major cause of morbidity and mortality in cancer patients and survivors.^{18,21} A structured international registry-level analysis of over 40 published case reports is summarized in Table 1, detailing cancer types, implicated therapies, timing of TCM onset, and diagnostic features (Table 1).

Table 1: An Expanded International Case Registry analysis table: Takotsubo Cardiomyopathy Associated with Cancer Drugs, which compiles over 40 case reports and studies.

Reference	Age	Sex	Cancer Type	Chemotherapy/ Immunotherapy	TCM Onset	ECG Abnormality	Peak Troponin (ng/ml)
29	60	F	Stage III Colorectal	5-FU (FOLFOX)	1st cycle	STE II, III, aVF, V4-V6	0.38
30	62	F	Rectal Adenocarcinoma	5-FU + Levofolinate	6th cycle	STE V1-V3	0.82
53	57	F	Renal cell carcinoma	Sunitinib	N/A	STE	1.2
48	50	F	Breast (HER2+)	Trastuzumab	11th cycle	TWI V1-2, I, aVL	0.15
6	58	F	Colon	FOLFOX	1st cycle	Non-specific ST changes	0.64
31	81	F	Colorectal	5-FU → Capecitabine	1st Capecitabine	TWI anteroseptal, inferior, lateral	0.35
87	55	M	AML	Cytarabine + Daunorubicin	1st cycle	STE I, aVL, V5, V6	38.64
62	83	F	Melanoma	Ipilimumab	4th cycle	STE I, V2, V3	0.98
54	71	F	Renal cell carcinoma	Axitinib	1st cycle	Anterolateral STE	6.95
66	73	F	DLBCL	LACE (Lomustine, Cytarabine, Cyclophosphamide, Etoposide)	Final day	TWI anterolateral	N/A
67	45	F	Anal Canal	Intra-arterial/IV polychemotherapy	1st cycle	STE	N/A
35	55	M	Caecum	Capecitabine	1st cycle	STE V2-6, STD aVR, II, aVF	89
51	54	F	Breast (HER2+)	Pertuzumab + Trastuzumab	2nd cycle	STE, TWI	0.21
56	71	F	Thyroid	Lenvatinib	First cycle	STE, TWI	0.9
65	66	M	Melanoma	Checkpoint Inhibitor	2nd cycle	STE, TWI	1.1
60	N/A	Various	Various	Immune checkpoint inhibitors	Various	STE, TWI, Sinus Tachycardia, etc.	0.2-2.5
28	62	F	Rectal adenocarcinoma	5-FU	1st cycle	STE, TWI	0.8
46	55	M	Caecum	Capecitabine	1st cycle	STE V2-6, STD aVR, II, aVF	89
5	62	F	Colon	5-FU	1st cycle	STE, TWI	0.7
88	62	F	Bladder	Gemcitabine + Cisplatin	2nd cycle	STE, TWI	0.6
81	50	F	Breast (HER2+)	Trastuzumab	2nd cycle	STE, TWI	0.5
49	48	F	Breast (HER2+)	Trastuzumab	1st cycle	STE, TWI	0.4
41	66/61	M	Colon/NSCLC	Bevacizumab	1st & 2nd cycle	STE, Q waves	N/A

82	2 cases (N/A)	F	Breast/ Colorectal	Bevacizumab/ Riltuximab+Chemo	1st cycle	STE, TWI	N/A
50	63	F	Breast (HER2+)	Pertuzumab + Trastuzumab + Nab-paclitaxel	3rd cycle	STE, TWI	N/A
83	55	F	Breast (HR+/HER2-)	Palbociclib + Letrozole	N/A	STE V2-V4	6.8
73	72	F	Colon/NSCLC	Regorafenib	10 days	STE, TWI	N/A
57	60s	F	NSCLC	Afatinib	N/A	STE, TWI	N/A
55	60s	F	Thyroid	Lenvatinib	N/A	STE, TWI	N/A
84	59	M	Lymphoma	RCHOP/RVD	Pre-chemo	STE, TWI	N/A
85	57	F	NSCLC	Carboplatin, Pemetrexed, Pembrolizumab	1st cycle	STE, TWI	N/A
69	60s	F	Lung	Gemcitabine + Vinorelbine	1st cycle	STE, TWI	N/A
68	54	F	Colon (recurrent)	CRS + HIPEC	Post-op	STE, TWI	N/A
88	60s	F	Palliative (various)	None (pain crisis)	N/A	STE, TWI	N/A
89	54	F	Rectosigmoid	Surgery	Post-op	TWI, QT prolongation	N/A
90	65	F	Ovarian (metastatic)	None (diagnosis)	N/A	STE, TWI	N/A
14	38	F	Endocervical	Chemoradiation + Brachytherapy	During brachy	STE, TWI	N/A
86	2 cases (N/A)	F	Breast/ Endometrial	Tamoxifen/ Radiotherapy	Adjuvant	STE, TWI	N/A
36	55	F	Breast (metastatic)	Capecitabine	1st cycle	STE, TWI	N/A
37	60s	F	Colorectal	Capecitabine	1st cycle	STE, TWI	N/A
47	50s	F	Colorectal	5-FU → Capecitabine rechallenge	After the prior 5-FU TCM	STE, TWI	N/A
44	39	F	Breast (advanced)	Capecitabine	During therapy	STE, TWI	N/A
52	50s	F	Breast (HER2+)	Trastuzumab	Early cycle	STE, TWI, LV thrombus	N/A
58	60s	M	CLL (ibrutinib therapy)	Ibrutinib	N/A	Mid-cavitary variant, STE	N/A
59	2 cases (N/A)	F	NSCLC (EGFR+)	EGFR-TKI (erlotinib/gefitinib)	Early cycles	STE, TWI	N/A
63	62	M	HCC	Nivolumab (ICI)	3 weeks after start	STE, arrhythmias	N/A
64	50	F	Cancer (unspecified combo therapy)	ICI + chemo	N/A	STE, TWI	N/A
45	59	M	Sigmoid colon	FOLFOX	1st cycle	STE, hyperacute T waves	1

5-fluorouracil (5-FU):

5-Fluorouracil is a fluoropyrimidine analog that inhibits thymidylate synthase, disrupting DNA synthesis and inducing apoptosis in malignant cells.^{1,22} It is administered by bolus or continuous infusion for solid tumors, most prominently colorectal, gastric, esophageal, and pancreatic cancers, and serves as a backbone in combination regimens such as FOLFOX and FOLFIRI with leucovorin modulation.^{23,24} Cardiotoxicity is well described, with overall cardiovascular event rates reported between 1.5% and 18% in clinical series, and spans vasospastic angina, arrhythmias, myocardial infarction, heart failure, and stress cardiomyopathy.^{5,25-27} Various case studies of cancer therapy-induced TCM are summarized in Table 1 (Table 1).

Multiple case reports have supported the association between 5-FU and Takotsubo cardiomyopathy. Kobayashi *et al.*

described a 62-year-old woman receiving first-cycle 5-FU for rectal adenocarcinoma who developed chest pain with ST-segment elevation; angiography revealed no obstructive coronary disease, and imaging demonstrated apical ballooning, consistent with TCM.²⁸ Basselin *et al.* reported a 60-year-old woman on first-cycle FOLFOX who presented with diffuse ST elevation and modest troponin rise, with ventriculography confirming apical dyskinesia;²⁹ the Naranjo adverse drug reaction scale indicated a probable relationship between 5-FU and TCM with a score of 8.^{1,29} Grunwald *et al.* described TCM after the sixth cycle of 5-FU plus leucovorin in rectal adenocarcinoma, showing that susceptibility can persist beyond initial exposure.³⁰ Ozturk *et al.* detailed an FOLFOX-treated patient who presented with an ACS-like picture and was ultimately diagnosed with TCM, underscoring the need for angiography and multimodality imaging to exclude obstructive CAD and myocarditis.⁶

Across reported cases, onset is most often during or immediately after continuous infusion, frequently in the first cycle, and the clinical pattern includes chest pain, ST-segment elevation or T-wave inversion, modest troponin elevation, unobstructed coronaries, and transient reduction in left ventricular ejection fraction with recovery within weeks after cessation.^{5,31,32} Severe complications have been described, including ventricular fibrillation and shock, although most patients recover with supportive care and guideline-based TCM management.¹ Mechanistically, 5-FU may provoke epicardial vasospasm in a subset of patients, supported by experimental and clinical observations, while others demonstrate microvascular dysfunction and catecholamine surges aligned with canonical TCM pathways.^{6,9} Broader toxicology series and pharmacovigilance reviews reinforce the plausibility of direct and indirect cardiotoxic mechanisms with fluoropyrimidines.³²⁻³⁴

Rechallenge after 5-FU-associated TCM is controversial. Abdelrahman *et al.* reported successful capecitabine rechallenge following 5-FU-induced TCM; however, other reviews describe high rates of recurrent cardiotoxicity and deaths with re-exposure to fluoropyrimidines, arguing for caution and individualized, multidisciplinary decision-making.³⁴⁻³⁷ When oncologic necessity is compelling, consideration of alternative regimens, inpatient monitoring, and preventive strategies for vasospasm may be appropriate, with strict adherence to TCM imaging and supportive therapy standards.^{10,38}

Taken together, the evidence places 5-FU among the most consistent chemotherapeutic triggers of TCM. The syndrome typically arises during early infusion, presents with ACS-like features in the absence of obstructive coronary disease, and resolves with drug cessation and supportive care. Because recurrence with rechallenge is not rare, decisions to re-expose should weigh oncologic benefit against cardiotoxic risk in a multidisciplinary framework.^{1,5}

Bevacizumab:

Bevacizumab is a recombinant humanized monoclonal antibody that targets vascular endothelial growth factor A (VEGF-A), thereby inhibiting angiogenesis and reducing tumor vascularization. By blocking VEGF signaling, bev-

acizumab enhances the efficacy of cytotoxic chemotherapy through decreased interstitial pressure and improved drug delivery to tumor tissue.^{18,39} Clinically, it is approved in combination regimens for metastatic colorectal cancer, non-small cell lung cancer, renal cell carcinoma, and several gynecologic malignancies.⁵⁴

The cardiovascular safety profile of bevacizumab has been a consistent concern in cardio-oncology. Hypertension is the most common adverse effect, but arterial thromboembolism, myocardial ischemia, heart failure, and arrhythmias have also been reported.^{21,40} Mechanistic studies suggest that VEGF inhibition disrupts endothelial nitric oxide signaling, reduces vascular repair capacity, and promotes microvascular dysfunction, all of which may predispose to myocardial injury and stress cardiomyopathy.³⁹ Inflammatory cytokine release and impaired coronary microcirculation have also been proposed as contributors to reversible left ventricular dysfunction in this setting.¹⁸

In sum, the evidence indicates that bevacizumab can act as a rare but plausible trigger for TCM in susceptible oncology patients. Proposed mechanisms include VEGF-related endothelial dysfunction, microvascular ischemia, and inflammatory signaling, all of which align with known pathways of stress cardiomyopathy. Although reported cases remain limited, clinicians should be alert to ACS-like presentations in patients receiving bevacizumab, especially those with existing cardiovascular risk factors. Prompt recognition, angiographic exclusion of obstructive disease, and supportive care are critical, as systolic function typically recovers once the drug is discontinued.³⁹⁻⁴¹

Capecitabine:

Capecitabine is an oral fluoropyrimidine prodrug that converts enzymatically to 5-fluorouracil (5-FU), with the final step catalyzed by thymidine phosphorylase, an enzyme concentrated in tumor tissue. This tumor-targeted activation was intended to improve tolerability compared to intravenous 5-FU while preserving antitumor activity.²³ Since its approval, capecitabine has been widely used in the treatment of colorectal, gastric, pancreatic, and breast cancers, either as monotherapy or in combination regimens such as CAPOX and XELOX.²⁵

However, clinical experience has demonstrated that capecitabine's cardiovascular toxicity profile is very similar to that of 5FU. Retrospective analyses report cardiotoxicity rates of approximately 3–9%, with manifestations ranging from angina and arrhythmias to myocardial infarction, cardiogenic shock, and sudden cardiac death.^{17,18,42} Symptoms typically develop within the first 72 hours of therapy and resolve after discontinuation, although severe complications have been described.

The most widely accepted mechanism is coronary vasospasm, supported by angiographic and clinical observations, but additional explanations include direct endothelial injury, prothrombotic effects, and immune-mediated myocardial inflammation.^{42,43} Experimental studies also suggest that fluoropyrimidines can impair microvascular function and mi-

tochondrial energetics, creating a substrate for stress-induced cardiomyopathy.³²

To date, seven cases of capecitabine-induced Takotsubo cardiomyopathy have been reported (Table 1). Molteni *et al.* described a 39-year-old woman with advanced breast cancer who developed acute chest pain and transient left ventricular dysfunction during capecitabine therapy, with full recovery after discontinuation of the drug.⁴⁴ Y-Hassan *et al.* reported a woman with advanced gastric cancer who presented with ST-segment elevation and classic wall-motion abnormalities, again resolving completely.⁴⁵ Endo *et al.* detailed a case complicated by both cardiomyopathy and multifocal cerebral leukoencephalopathy during capecitabine therapy.⁴⁶ Qasem *et al.* described a metastatic breast cancer patient who developed TCM during the first cycle of treatment, with unobstructed coronaries and normalization of function within weeks.³⁶ Abdelrahman *et al.* documented a rechallenge scenario in which capecitabine was successfully reintroduced after prior 5-FU-induced TCM under close monitoring.⁴⁷ Bhardwaj *et al.* reported a 47-year-old woman with metastatic breast cancer who developed chest pain, ST-elevation, and apical ballooning during her first cycle, with full recovery.³⁵ Abdelmaseih *et al.* presented a colorectal cancer patient with ACS-like features who was ultimately diagnosed with TCM and recovered after cessation.³⁷ Across these reports, patients ranged from middle-aged to elderly, had no prior coronary artery disease, typically developed symptoms during the first cycle, and all regained left ventricular function within 1–4 weeks.

Across these reports, consistent features emerge: patients were between 47 and 81 years of age, without prior coronary artery disease, and most developed symptoms during the first cycle of therapy. Presentations included chest pain, ECG changes with modest troponin elevation, and unobstructed coronaries on angiography. All patients demonstrated transient left ventricular systolic dysfunction with recovery of ejection fraction within 1–4 weeks after cessation of therapy.

Taken together, these findings indicate that capecitabine, like 5-FU, can act as a rare but credible trigger of Takotsubo cardiomyopathy. Clinicians should maintain vigilance for ACS-like presentations in patients receiving capecitabine, particularly during early cycles, and weigh the risks of rechallenge against oncologic necessity in a multidisciplinary framework.

Trastuzumab:

Trastuzumab is a recombinant humanized monoclonal antibody targeting the extracellular domain of the HER2 receptor, a member of the epidermal growth factor receptor family. By blocking HER2 signaling, trastuzumab inhibits tumor cell proliferation and survival, and it has become a cornerstone of therapy for HER2-positive breast and gastric cancers.¹⁷ Its introduction dramatically improved outcomes in these malignancies, but it also established a new paradigm of targeted therapy-related cardiotoxicity.

The most recognized cardiac complication of trastuzumab is left ventricular systolic dysfunction and heart failure, often reversible with drug discontinuation and standard therapy.³⁸ Unlike anthracyclines, trastuzumab does not cause cumulative

dose-dependent myocyte necrosis; instead, it interferes with HER2-mediated survival pathways in cardiomyocytes, impairing stress adaptation and repair.³⁹ This mechanism provides a biologic rationale for reversible dysfunction, but it also suggests susceptibility to stress-induced syndromes such as Takotsubo cardiomyopathy (TCM).

Several published cases have directly linked trastuzumab with TCM. Khanji *et al.* described a patient who developed acute chest pain and ST-segment elevation after trastuzumab infusion, with unobstructed coronaries and apical ballooning consistent with TCM.⁴⁸ Burgy *et al.* reported a woman who had postoperative TCM and was later re-treated with trastuzumab without recurrence, highlighting the complexity of re-exposure decisions.⁴⁹ Lees *et al.* presented a case of TCM during anti-HER2 therapy for metastatic breast cancer, again with full recovery of ventricular function. Matsumoto *et al.* reported a case of TCM triggered by an infusion reaction to trastuzumab, pointing to the impact of acute stress responses.⁵⁰ Patruno *et al.* described TCM during dual HER2 blockade with pertuzumab and trastuzumab, raising concern about increased risk.⁵¹ Patel *et al.* most recently described a complex case of trastuzumab-induced TCM with left ventricular thrombus and embolic stroke, showing that although recovery is common, major problems can happen.⁵²

The majority of patients across these reports were female and had HER2-positive breast cancer, which usually manifested during or soon after trastuzumab infusion. Chest discomfort, ST-segment elevation, mild troponin increase, unobstructed coronaries, and temporary apical ballooning were among the clinical characteristics. After trastuzumab was stopped, left ventricular function recovered in almost all cases within a few weeks, although there have been a few documented unusual side effects, including thrombus formation.

Overall, the data suggest that trastuzumab may be a rare cause of Takotsubo cardiomyopathy, even though it is mainly linked to reversible LV dysfunction. The suggested mechanism combines acute stress reactions during infusion with impairment of HER2-mediated cardioprotective signaling. Rechallenge decisions should be made on an individual basis within a multidisciplinary cardio-oncology framework, and clinicians should remain vigilant for ACS-like symptoms in patients receiving trastuzumab, especially early in therapy.

Tyrosine Kinase Inhibitors (TKIs):

Tyrosine kinase inhibitors (TKIs) are small-molecule agents that target intracellular signaling pathways essential for tumor growth and angiogenesis. By inhibiting receptors such as VEGFR, PDGFR, EGFR, and RET, TKIs have become central to the treatment of renal cell carcinoma, thyroid cancer, and non-small cell lung cancer. Their cardiovascular safety profile, however, has raised concern, with hypertension, QT prolongation, heart failure, and ischemic events well documented.^{39,40} More recently, several TKIs have been implicated in TTS, suggesting that endothelial dysfunction, microvascular ischemia, and acute hypertensive stress may converge with canonical catecholamine-mediated pathways to trigger this syndrome.

Sunitinib was the first TKI linked to TCM, with Numico *et al.* describing a renal cancer patient who developed acute chest pain and ST-segment elevation during therapy; angiography revealed unobstructed coronaries, and ventricular function normalized after drug withdrawal.⁵³ *Axitinib* has also been implicated, with Ovadia *et al.* reporting a patient who developed apical ballooning and transient systolic dysfunction consistent with TCM, again resolving after cessation.⁵⁴ *Lenvatinib* has been associated with multiple cases: Chae *et al.* described a thyroid cancer patient who developed both posterior reversible encephalopathy syndrome and TCM while on lenvatinib,⁵⁵ while Fashafsha *et al.* reported another thyroid cancer case in which lenvatinib triggered TCM, with recovery after discontinuation.⁵⁶ *Afatinib* has similarly been implicated, with Ramos *et al.* presenting a non-small cell lung cancer patient who developed TCM during therapy, characterized by apical ballooning and normalization of function within weeks.⁵⁷ Beyond VEGF and EGFR inhibitors, Giza *et al.* reported a mid-cavitary variant of TCM in a patient receiving the Bruton's tyrosine kinase inhibitor *ibrutinib*, expanding the spectrum of chemotherapy-associated TCM beyond the typical apical form.⁵⁸ Most recently, Tanaka *et al.* described two additional cases of TCM during EGFR-TKI therapy for non-small cell lung cancer with EGFR mutations, underscoring that this may represent a broader class effect.⁵⁹

Across these reports, patients typically presented with chest pain, ST-segment elevation, modest troponin rise, and unobstructed coronaries, with echocardiography or ventriculography confirming transient wall-motion abnormalities. Most cases occurred early in therapy, often within the first cycles, and all demonstrated recovery of left ventricular function within 1–4 weeks after cessation.

Taken together, these findings indicate that TKIs, though mechanistically distinct from fluoropyrimidines and monoclonal antibodies, can act as rare but credible triggers of Takotsubo cardiomyopathy. The proposed mechanisms include VEGF-mediated endothelial dysfunction, acute hypertensive stress, and microvascular ischemia, which align with canonical catecholamine-driven pathways. Clinicians should maintain vigilance for ACS-like presentations in patients receiving TKIs, particularly during early cycles, and weigh oncologic benefit against cardiotoxic risk in a multidisciplinary framework.

Immune Checkpoint Inhibitors (ICIs):

Immune Checkpoint Inhibitors (ICIs) Immune checkpoint inhibitors (ICIs) such as *ipilimumab*, *nivolumab*, *pembrolizumab*, and *atezolizumab* have revolutionized the treatment of melanoma, lung cancer, and other solid tumors by unleashing antitumor immunity through blockade of *CTLA-4* or *PD-1/PD-L1* pathways. While their most recognized cardiovascular toxicities include myocarditis, pericarditis, and arrhythmias, a growing body of evidence implicates ICIs in TCM. The proposed mechanisms include immune-mediated myocardial inflammation, cytokine surge, and catecholamine excess triggered by systemic immune activation, which converge with canonical stress-cardiomyopathy pathways.^{60,61}

The first published case was by Geisler *et al.*, who described a melanoma patient treated with ipilimumab who developed acute chest pain and apical ballooning consistent with TCM, with full recovery after discontinuation.⁶² Tan *et al.* reported a 62-year-old man with hepatocellular carcinoma who developed concomitant myopericarditis and TCM three weeks after nivolumab, presenting with ST-segment elevation and ventricular arrhythmias; coronary angiography was normal, and ventricular function normalized with supportive care.⁶³ Yamada *et al.* described a cancer patient receiving combination therapy, including an ICI, who developed ACS-like symptoms and was ultimately diagnosed with TCM, underscoring the diagnostic challenge in polytherapy settings.⁶⁴ Serzan *et al.* presented a series of ICI-associated TCM cases, including patients treated with nivolumab and pembrolizumab, all of whom demonstrated transient LV dysfunction with recovery after drug withdrawal.⁶⁵ More recently, systematic reviews by Kyriakoulis *et al.* and Trontzas *et al.* synthesized all published cases to date, confirming that TCM has been reported across multiple ICIs, including ipilimumab, nivolumab, pembrolizumab, atezolizumab, avelumab, and durvalumab.⁶⁰

Across these reports, patients typically presented with chest pain, ST-segment elevation or T-wave inversion, modest troponin rise, and unobstructed coronaries, with echocardiography or ventriculography confirming transient wall-motion abnormalities. Most cases occurred within the first few cycles of therapy, and all demonstrated recovery of left ventricular function within weeks after cessation. In sum, these findings indicate that ICIs, while primarily associated with myocarditis, can also act as rare triggers of Takotsubo cardiomyopathy. The overlap of immune-mediated inflammation and catecholamine-driven stress responses likely explains this association. Clinicians should maintain vigilance for ACS-like presentations in patients receiving ICIs, particularly early in therapy, and decisions regarding rechallenge should be individualized within a multidisciplinary cardio-oncology framework.

Polychemotherapy Regimens:

Takotsubo cardiomyopathy has also been described in the context of multi-agent chemotherapy, where attribution to a single drug is not possible. Malley *et al.* reported a patient who developed acute chest pain and transient apical ballooning after *combination chemotherapy*, with full recovery following supportive care.⁶⁶ Similarly, Coen *et al.* described TCM in a patient treated with a polychemotherapy regimen for anal canal carcinoma, and their accompanying review emphasized that heterogeneous regimens, not only fluoropyrimidines, can precipitate the syndrome.⁶⁷ These cases highlight the cumulative cardiotoxic stress that may arise from multi-drug exposure.

Other Oncologic Triggers:

Beyond systemic therapy, oncologic interventions themselves have been implicated in TCM. Vieira *et al.* reported a case of ventricular fibrillation arrest and subsequent TCM during *brachytherapy* for endocervical adenocarcinoma, with unobstructed coronaries on angiography.¹⁴ Moldovan *et al.* described life-threatening TCM following cytoreductive sur-

gery with *hyperthermic intraperitoneal chemotherapy (HIPEC)* for recurrent colon cancer, underscoring perioperative stress as a potent trigger.⁶⁸ These reports illustrate that procedural and surgical stressors in oncology can act as direct precipitants of stress cardiomyopathy.

Miscellaneous Single-Agent Reports:

Several isolated case reports expand the spectrum of agents associated with TCM. Özbay *et al.* described TCM occurring shortly after administration of “non-cardiotoxic” chemotherapy with *vinorelbine* and *gemcitabine*, reinforcing that stress cardiomyopathy can follow agents without classic cardiotoxic profiles.⁶⁹ Additional scattered reports have implicated *oxaliplatin* and *cyclophosphamide* in ACS-like presentations ultimately diagnosed as TCM, though these remain rare and anecdotal.^{27,42} Collectively, these cases suggest that while certain drug classes carry stronger mechanistic links, virtually any oncologic therapy may act as a trigger in susceptible patients.

Mechanistic Pathways of Takotsubo Cardiomyopathy in Oncology (Pathophysiology):

Catecholamine Surge and Sympathetic Overdrive:

Microvascular dysfunction is another central mechanism. Imaging and invasive studies have shown impaired coronary flow reserve and abnormal microvascular reactivity in TCM patients despite angiographically normal epicardial arteries.⁷⁰ Endothelial dysfunction and impaired nitric oxide signaling contribute to transient ischemia at the tissue level, which aligns with the reversible nature of the syndrome. This mechanism is particularly relevant in oncology, as VEGF inhibitors and fluoropyrimidines are known to impair microvascular function and may exacerbate susceptibility to stress cardiomyopathy.¹

Epicardial Coronary Spasm:

Although less common, epicardial coronary spasm has been documented in a subset of patients. Early Japanese reports described multivessel spasm during angiography in TCM, suggesting that vasospasm can mimic or precipitate the syndrome.^{3,7} This mechanism is especially relevant in fluoropyrimidine-associated cases, where vasospasm is a well-recognized cardiotoxic effect.

Inflammatory and Immune Pathways:

Inflammatory activation also appears to play a role. Cardiac magnetic resonance often demonstrates myocardial edema without necrosis, consistent with cytokine-mediated injury.⁸ Immune checkpoint inhibitors provide a clinical link, as several cases of TCM have occurred in parallel with myocarditis or systemic immune activation, suggesting overlapping inflammatory and catecholamine-driven pathways.⁶⁰ Singh *et al.* emphasized that neurohormonal stress and immune dysregulation likely act in concert, producing a vulnerable substrate for transient dysfunction.⁷¹

Sex and Age Predisposition in Takotsubo Cardiomyopathy:

Epidemiologic studies consistently demonstrate that Takotsubo cardiomyopathy occurs predominantly in women, with

over 90% of cases reported in postmenopausal patients.⁷ The InterTAK registry confirmed this striking demographic pattern, noting that the typical patient is a woman in her sixth or seventh decade of life, often presenting after an identifiable emotional or physical stressor.⁸ Several mechanisms have been proposed to explain this predisposition. Loss of estrogen after menopause is thought to reduce cardioprotective signaling, impair endothelial function, and increase susceptibility to catecholamine-mediated injury.⁷² Experimental and clinical data also suggest that postmenopausal women exhibit altered β -adrenergic receptor responsiveness and greater coronary microvascular dysfunction, both of which may amplify the myocardial effects of acute stress.^{9,71}

In the oncology setting, this demographic vulnerability is particularly relevant. Cancer therapies that provoke sympathetic activation, endothelial stress, or inflammatory signaling may act on an already susceptible substrate, explaining why many reported cases of chemotherapy-associated TCM involve middle-aged or older women. Recognizing this pattern is clinically important, as it should heighten suspicion for TCM in postmenopausal women presenting with ACS-like symptoms during cancer treatment.^{7,8}

Oncologic Context:

Cancer itself may heighten susceptibility to TCM through chronic stress, systemic inflammation, and paraneoplastic mediators. Desai *et al.* highlighted that malignancy is over-represented in TCM cohorts, with prevalence estimates of 10–17% in modern registries.¹ Paraneoplastic cytokines and altered adrenergic receptor signaling may further sensitize the myocardium to stress. When combined with cardiotoxic therapies, these factors create a “two-hit” model in which both the underlying cancer and its treatment converge on the same catecholamine-microvascular axis.^{70,71}

Standard/Conventional Treatments and Prognosis:

The management of Takotsubo cardiomyopathy remains largely supportive, guided by principles of acute heart failure care and tailored to the hemodynamic profile of each patient. Initial evaluation requires exclusion of acute coronary syndromes through urgent angiography and multimodality imaging, with echocardiography and cardiac magnetic resonance playing central roles in confirming the diagnosis and identifying complications such as thrombus or myocarditis.^{8,10} Once obstructive coronary disease is ruled out, therapy is directed toward stabilizing ventricular function. Patients with reduced ejection fraction are typically treated with beta-blockers, ACE inhibitors or ARBs, and diuretics as needed, with continuation until recovery of systolic function, which usually occurs within weeks.^{8,72,75}

Complications require specific management strategies. In cases of left ventricular outflow tract obstruction, inotropes should be avoided, as catecholamine stimulation can worsen obstruction; instead, cautious beta-blockade and volume expansion are recommended.^{72,73,75} When the left ventricular thrombus is present, or the ejection fraction is severely depressed, anticoagulation is indicated until resolution is doc-

umented.⁸ Arrhythmias and QT prolongation are managed with telemetry, electrolyte optimization, and avoidance of QT-prolonging drugs.⁸ In patients with cardiogenic shock, vasopressors may be required, but catecholamine agents should be minimized; mechanical circulatory support such as ECMO or percutaneous ventricular assist devices can be considered in refractory cases, while intra-aortic balloon pump use is controversial in the presence of outflow obstruction.^{8,72,75}

Prognosis is generally favorable, with most patients achieving full recovery of ventricular function within four to eight weeks, as confirmed in large registry studies.⁷ Nevertheless, recurrence is possible, and long-term follow-up with echocardiography is advised. Perioperative and anesthetic management also requires vigilance, as acute stressors can precipitate recurrence; careful hemodynamic monitoring and avoidance of excessive sympathetic stimulation are recommended in surgical settings.¹³ Taken together, conventional therapy for TCM emphasizes supportive care, complication-specific strategies, and close follow-up, with outcomes that are typically excellent when the condition is promptly recognized and managed.

Ongoing Clinical Studies on The Mechanism of Novel Therapeutic Interventions in TCM:

Several ongoing and completed clinical studies provide context for Takotsubo cardiomyopathy (TCM) in oncology. The PSYCHONIC study (NCT03437642) explores psychosomatic determinants of stress cardiomyopathy, profiling psychological stressors and autonomic markers relevant to cancer patients.⁷⁰ The TTAC registry (NCT01944826) systematically captures the timing of cancer diagnosis and treatment relative to TCM onset, lending epidemiologic weight to the observed overrepresentation of malignancy in TCM cohorts.⁷⁴ The MAESTRO trial (NCT05793008) prospectively distinguishes primary versus secondary TCM, a classification particularly applicable to oncology, where therapy-related triggers dominate.⁷⁶ Complementing these, the Psychosocial Support study (NCT05977049) evaluates behavioral interventions to reduce anxiety in TTS survivors, offering a framework for supportive care that could be adapted to cancer populations.⁷⁷ Finally, the EACVI-INFLAME study (NCT07077304) standardizes multimodality imaging across inflammatory cardiovascular diseases, including TCM, directly addressing diagnostic rigor and complication tracking.⁷⁸ Together, these trials underscore the multifactorial nature of cancer-associated TCM and highlight the need for integrated psychosocial, epidemiologic, and imaging approaches to improve recognition and management (Table 2).

Table 2: Summary of selected clinical trials investigating Takotsubo cardiomyopathy.

Clinical Trials.gov ID	Design	Population	Primary endpoint
NCT03437642	Observational	Oncologic and cardiac patients with suspected or confirmed TCM	Association of psychosocial/autonomic profiles with TCM onset and outcomes
NCT01944826	Registry (observational)	TCM patients with and without malignancy	Prevalence/temporal association of cancer and therapy with TCM; imaging and short-term outcomes
NCT05793008	Prospective characterization	Adults with primary or secondary TTS	Phenotypic classification with imaging/biomarkers; clinical outcomes
NCT05977049	Interventional (behavioral)	Adults with prior TTS	Anxiety reduction at 12 months with a structured support program
NCT07077304	Observational	Patients with inflammatory CVD, including TCM	Diagnostic standardization and validation across echo/CMR/CT/nuclear for TCM and related conditions

■ **Results and Discussion**

Takotsubo cardiomyopathy (TCM) has progressed from a rare syndrome first described in Japan to a recognized clinical entity with distinct epidemiology, pathophysiology, and outcomes. Large registries such as InterTAK established their demographic profile, predominantly postmenopausal women, and highlighted the role of acute stressors and neurohumoral activation in its pathogenesis.^{7,8} Mechanistic studies confirmed catecholamine excess, microvascular dysfunction, and, in select cases, epicardial spasm as central drivers of transient myocardial stunning.^{9,70,73} These insights provided the foundation for exploring how cancer and its therapies intersect with TCM, given the profound physical, emotional, and inflammatory stressors inherent to oncology care.^{1,4}

This review demonstrates that malignancy is not only overrepresented in TCM cohorts but also shapes the clinical phenotype. Cancer patients more often present with secondary TCM triggered by illness or therapy, and their short-term outcomes appear worse than those of emotionally triggered cases.^{1,16} Case reports consistently implicate fluoropyrimidines, particularly 5-fluorouracil and capecitabine, as reproducible pharmacologic triggers, with mechanistic plausibility supported by vasospasm, microvascular dysfunction, and catecholamine surges.^{6,29,30} Targeted therapies such as trastuzumab and bevacizumab, as well as tyrosine kinase inhibitors and immune checkpoint inhibitors, have also been linked to TCM, underscoring the convergence of endothelial injury, adrenergic stress, and immune activation on a vulnerable myocardium.^{41,48,60,65} These drug-specific observations align with the broader mechanistic framework of TCM and reinforce the need for vigilance in oncology practice.

A major clinical dilemma is whether to discontinue or re-challenge cancer therapy after a TCM episode. While case reports and pharmacovigilance data support drug-induced TCM, many lack complete diagnostic workups, raising the

possibility that some events reflect stress cardiomyopathy from the cancer itself or its psychosocial burden.^{1,6} Prematurely discontinuing an effective anti-cancer agent can compromise oncologic outcomes, particularly when the drug is central to curative regimens.² Conversely, rechallenge carries a real risk of recurrence, as documented with fluoropyrimidines and other agents.^{37,47} This underscores the need for individualized, multidisciplinary decision-making that balances oncologic benefit against cardiotoxic risk, ideally supported by prospective data rather than isolated case reports.

The clinical studies summarized in Table 2 extend these observations by moving beyond anecdotal reports toward structured investigation. The TTAC registry (NCT01944826) and PSYCHONIC study (NCT03437642) explicitly address the cancer-TCM intersection, capturing epidemiology and psychosocial determinants. MAESTRO (NCT05793008) advances phenotypic classification by distinguishing primary from secondary TCM, a distinction highly relevant to oncology patients. The psychosocial support trial (NCT05977049) operationalizes the brain-heart axis by testing whether structured interventions can mitigate recurrence and improve quality of life. Finally, EACVI-INFLAME (NCT07077304) emphasizes diagnostic rigor through standardized multimodality imaging, addressing the persistent challenge of differentiating TCM from ACS or myocarditis. Together, these studies illustrate how the field is transitioning from descriptive case reports to prospective, mechanistic, and interventional research.

Despite these advances, important limitations remain. Much of the evidence linking chemotherapy and targeted therapies to TCM derives from case reports and small series, many of which lack complete diagnostic workups such as cardiac MRI or biopsy.^{1,6} Registries provide valuable epidemiologic insights but are often limited by referral bias and incomplete oncologic data. Clinical trials remain few, and several have unknown or inactive status, underscoring the need for sustained prospective enrollment.

Looking forward, future research should prioritize three directions. First, prospective registries that integrate oncology and cardiology data are needed to clarify the true incidence, timing, and outcomes of cancer-associated TCM. Second, mechanistic studies should explore how specific therapies—fluoropyrimidines, HER2-targeted agents, VEGF inhibitors, and immune checkpoint inhibitors—interact with adrenergic, endothelial, and immune pathways to precipitate TCM. Third, interventional trials are required to test preventive and therapeutic strategies, including pharmacologic prophylaxis, structured psychosocial support, and cardio-oncology monitoring frameworks. By combining mechanistic insight with prospective clinical data, the field can move toward evidence-based strategies that balance oncologic efficacy with cardiovascular safety.

■ Conclusion

Takotsubo cardiomyopathy represents a unique intersection of oncology and cardiology, where the physiological and psychological stresses of cancer, combined with the car-

diotoxic potential of modern therapies, create a heightened vulnerability to this reversible yet serious syndrome. Although most patients recover ventricular function, the occurrence of TCM in the setting of malignancy is associated with worse short-term outcomes and, critically, the potential interruption of life-saving cancer therapy. This review underscores that distinguishing true drug-induced TCM from other reversible cardiotoxicities remains a central challenge, as premature discontinuation of effective regimens can compromise oncologic survival, while rechallenge carries a risk of recurrence. Moving forward, careful case-by-case evaluation, integration of multimodality imaging, and prospective research into mechanisms and preventive strategies are essential to ensure that patients receive both optimal cancer treatment and vigilant cardiovascular protection.

■ Acknowledgments

I would like to sincerely thank Professor Prahlad Parajuli of Wayne State University for his mentorship and guidance throughout the development of this review. His expertise in tumor immunology and cardio-oncology helped me frame Takotsubo cardiomyopathy within the broader context of cancer-related stress and therapy-induced cardiotoxicity. Professor Parajuli encouraged me to approach the literature with rigor, emphasizing systematic extraction of case data and careful interpretation of mechanisms, while also reminding me to balance clinical detail with translational relevance. His feedback on methodology and diagnostic considerations strengthened the clarity and impact of this work. This study received no external funding and was conducted independently using publicly available sources.

■ References

- Desai, A.; Noor, A.; Joshi, S.; Kim, A. S. Takotsubo Cardiomyopathy in Cancer Patients. *Cardio-Oncology* **2019**, *5* (1). <https://doi.org/10.1186/s40959-019-0042-9>.
- Madias, J. E. Is Takotsubo Syndrome in Patients Receiving Chemotherapy Drug-Specific? *World Journal of Clinical Cases* **2015**, *3* (2), 204. <https://doi.org/10.12998/wjcc.v3.i2.204>.
- Dote, K.; Sato, H.; Tateishi, H.; Uchida, T.; Ishihara, M. [Myocardial Stunning due to Simultaneous Multivessel Coronary Spasms: A Review of 5 Cases]. *Journal of Cardiology* **1991**, *21* (2), 203–214.
- Tini, G.; Arcari, L.; Raffaella Mistrulli; Follesa, F.; Cianca, A.; Sclafani, M.; Tocci, G.; Paolo Spallarossa; Battistoni, A.; Cacciotti, L.; Musumeci, B.; Barbato, E. A Contemporary Update on Cancer and Takotsubo Syndrome. *Frontiers in cardiovascular medicine* **2024**, *10*. <https://doi.org/10.3389/fcvm.2023.1301383>.
- Stephanie Hui-Su Lim; Sharon Mary Wilson; Hunter, A.; Hill, J.; Beale, P. Takotsubo Cardiomyopathy and 5-Fluorouracil: Getting to the Heart of the Matter. *Case Reports in Oncological Medicine* **2013**, *2013*, 1–4. <https://doi.org/10.1155/2013/206765>.
- Ozturk, M. A.; Ozveren, O.; Cinar, V.; Erdik, B.; Oyan, B. Takotsubo Syndrome. *Blood Coagulation & Fibrinolysis* **2013**, *24* (1), 90–94. <https://doi.org/10.1097/mbc.0b013e3283597605>.
- Templin, C.; Ghadri, J. R.; Diekmann, J.; Napp, L. C.; Bataiosu, D. R.; Jaguszewski, M.; Cammann, V. L.; Sarcon, A.; Geyer, V.; Neumann, C. A.; Seifert, B.; Hellermann, J.; Schwyzer, M.; Eisenhardt, K.; Jenewein, J.; Franke, J.; Katus, H. A.; Burgdorf, C.; Schunkert, H.; Moeller, C. Clinical Features and Outcomes of Takotsubo

- (Stress) Cardiomyopathy. *New England Journal of Medicine* **2015**, 373 (10), 929–938. <https://doi.org/10.1056/nejmoa1406761>.
8. Ghadri, J.-R.; Wittstein, I. S.; Prasad, A.; Sharkey, S.; Dote, K.; Akashi, Y. J.; Cammann, V. L.; Crea, F.; Galiuto, L.; Desmet, W.; Yoshida, T.; Manfredini, R.; Eitel, I.; Kosuge, M.; Nef, H. M.; Deshmukh, A.; Lerman, A.; Bossone, E.; Citro, R.; Ueyama, T. International Expert Consensus Document on Takotsubo Syndrome (Part I): Clinical Characteristics, Diagnostic Criteria, and Pathophysiology. *European Heart Journal* **2018**, 39 (22), 2032–2046. <https://doi.org/10.1093/eurheartj/ehy076>.
 9. Wittstein, I. S.; Thiemann, D. R.; Lima, J. A. C.; Baughman, K. L.; Schulman, S. P.; Gerstenblith, G.; Wu, K. C.; Rade, J. J.; Bivalacqua, T. J.; Champion, H. C. Neurohumoral Features of Myocardial Stunning due to Sudden Emotional Stress. *New England Journal of Medicine* **2005**, 352 (6), 539–548. <https://doi.org/10.1056/nejmoa043046>.
 10. Citro, R.; Okura, H.; Ghadri, J. R.; Izumi, C.; Meimoun, P.; Izumo, M.; Dawson, D.; Kaji, S.; Eitel, I.; Kagiya, N.; Kobayashi, Y.; Templin, C.; Delgado, V.; Nakatani, S.; Popescu, B. A.; Bertrand, P.; Donal, E.; Dweck, M.; Galderisi, M.; Haugaa, K. H. Multimodality Imaging in Takotsubo Syndrome: A Joint Consensus Document of the European Association of Cardiovascular Imaging (EACVI) and the Japanese Society of Echocardiography (JSE). *European Heart Journal - Cardiovascular Imaging* **2020**, 21 (11), 1184–1207. <https://doi.org/10.1093/ehjci/jeaa149>.
 11. Madhavan, M.; Prasad, A. Proposed Mayo Clinic Criteria for the Diagnosis of Tako-Tsubo Cardiomyopathy and Long-Term Prognosis. *Herz* **2010**, 35 (4), 240–244. <https://doi.org/10.1007/s00059-010-3339-x>.
 12. Díaz-Navarro, R. Takotsubo Syndrome: The Broken-Heart Syndrome. *British Journal of Cardiology* **2021**. <https://doi.org/10.5837/bjc.2021.011>.
 13. Pillitteri, M.; Brogi, E.; Piagnani, C.; Bozzetti, G.; Forfori, F. Perioperative Management of Takotsubo Cardiomyopathy: An Overview. *Journal of Anesthesia, Analgesia and Critical Care* **2024**, 4 (1). <https://doi.org/10.1186/s44158-024-00178-y>.
 14. Vieira, A. C.; Ribeiro, M. F. S. A.; Lima, J.; Filho, J. S.; de Andrade Carvalho, H.; Mano, M. S. Takotsubo Syndrome Induced by Brachytherapy in a Patient with Endocervical Adenocarcinoma. *Cardio-oncology (London, England)* **2020**, 6 (1), 30. <https://doi.org/10.1186/s40959-020-00082-8>.
 15. Cameli, M.; Landra, F. Editorial: Case Reports in Heart Failure and Transplantation: 2022. *Frontiers in Cardiovascular Medicine* **2023**, 10. <https://doi.org/10.3389/fcvm.2023.1282064>.
 16. Cammann, V. L.; Annahita Sarcon; Ding, K. J.; Seifert, B.; Kato, K.; Davide Di Vece; Szawan, K. A.; Gili, S.; Jurisic, S.; Bacchi, B.; Micek, J.; Frangieh, A. H.; L. Christian Napp; Miłosz Jaguszewski; Bossone, E.; Citro, R.; Fabrizio D'Ascenzo; Franke, J.; Noutsias, M.; Knorr, M. Clinical Features and Outcomes of Patients with Malignancy and Takotsubo Syndrome: Observations from the International Takotsubo Registry. *Journal of the American Heart Association* **2019**, 8 (15). <https://doi.org/10.1161/jaha.118.010881>.
 17. Trapani, D.; Zagami, P.; Nicolò, E.; Pravettoni, G.; Curigliano, G. Management of Cardiac Toxicity Induced by Chemotherapy. *Journal of Clinical Medicine* **2020**, 9 (9), 2885. <https://doi.org/10.3390/jcm9092885>.
 18. Barachini, S.; Ghelardoni, S.; Varga, Z. V.; Mehanna, R. A.; Montt-Guevara, M. M.; Ferdinandy, P.; Madonna, R. Antineoplastic Drugs Inducing Cardiac and Vascular Toxicity – an Update. *Vascular Pharmacology* **2023**, 153, 107223. <https://doi.org/10.1016/j.vph.2023.107223>.
 19. Han, X.; Zhou, Y.; Liu, W. Precision Cardio-Oncology: Understanding the Cardiotoxicity of Cancer Therapy. *npj Precision Oncology* **2017**, 1 (1). <https://doi.org/10.1038/s41698-017-0034-x>.
 20. Peng, J.; Rushton, M.; Johnson, C.; Brezden-Masley, C.; Sulpher, J.; Chiu, M. G.; Graham, I. D.; Dent, S. An International Survey of Healthcare Providers' Knowledge of Cardiac Complications of Cancer Treatments. *Cardio-Oncology* **2019**, 5 (1). <https://doi.org/10.1186/s40959-019-0049-2>.
 21. de Wall, C.; Bauersachs, J.; Berliner, D. Cardiooncology—Dealing with Modern Drug Treatment, Long-Term Complications, and Cancer Survivorship. *Clinical & Experimental Metastasis* **2021**, 38 (4), 361–371. <https://doi.org/10.1007/s10585-021-10106-x>.
 22. Cobarro, L.; Valbuena-López, S.; Contreras, C.; Ugue-to, C.; Ruiz-Gutiérrez, I.; Peña-López, J.; Ruiz-Giménez, L.; López-Fernández, T. Immune Checkpoint Inhibitor-Related Stress Cardiomyopathy. *JACC: Case Reports* **2023**, 16, 101881. <https://doi.org/10.1016/j.jaccas.2023.101881>.
 23. Parker, W. B.; Cheng, Y. C. Metabolism and Mechanism of Action of 5-Fluorouracil. *Pharmacology & Therapeutics* **1990**, 48 (3), 381–395. [https://doi.org/10.1016/0163-7258\(90\)90056-8](https://doi.org/10.1016/0163-7258(90)90056-8).
 24. Goldberg, R. M.; Tabah-Fisch, I.; Bleiberg, H.; de Gramont, A.; Tournigand, C.; Andre, T.; Rothenberg, M. L.; Green, E.; Sargent, D. J. Pooled Analysis of Safety and Efficacy of Oxaliplatin plus Fluorouracil/Leucovorin Administered Bimonthly in Elderly Patients with Colorectal Cancer. *Journal of Clinical Oncology* **2006**, 24 (25), 4085–4091. <https://doi.org/10.1200/jco.2006.06.9039>.
 25. Heinemann, V.; von Weikersthal, L. F.; Decker, T.; Kiani, A.; Vehlting-Kaiser, U.; Al-Batran, S.-E.; Heintges, T.; Lerchenmüller, C.; Kahl, C.; Seipelt, G.; Kullmann, F.; Stauch, M.; Scheithauer, W.; Hielscher, J.; Scholz, M.; Müller, S.; Link, H.; Niederle, N.; Rost, A.; Höffkes, H.-G. FOLFIRI plus Cetuximab versus FOLFIRI plus Bevacizumab as First-Line Treatment for Patients with Metastatic Colorectal Cancer (FIRE-3): A Randomised, Open-Label, Phase 3 Trial. *The Lancet Oncology* **2014**, 15 (10), 1065–1075. [https://doi.org/10.1016/s1470-2045\(14\)70330-4](https://doi.org/10.1016/s1470-2045(14)70330-4).
 26. Labianca, R.; Beretta, G.; Clerici, M.; Frascini, P.; Luporini, G. Cardiac Toxicity of 5-Fluorouracil: A Study on 1083 Patients. *Tumori Journal* **1982**, 68 (6), 505–510. <https://doi.org/10.1177/030089168206800609>.
 27. Dalzell, J. R.; Samuel, L. M. The Spectrum of 5-Fluorouracil Cardiotoxicity. *Anti-Cancer Drugs* **2008**, 20 (1), 79–80. <https://doi.org/10.1097/cad.0b013e3283165f27>.
 28. Kobayashi, N.; Hata, N.; Yokoyama, S.; Takuro Shinada; Akihiro Shirakabe; Mizuno, K. A Case of Takotsubo Cardiomyopathy during 5-Fluorouracil Treatment for Rectal Adenocarcinoma. *Journal of Nippon Medical School* **2009**, 76 (1), 27–33. <https://doi.org/10.1272/jnms.76.27>.
 29. P D Cécile Basselin; T. Fontanges; Descotes, J.; Chevalier, P.; Bui-Xuan, B.; P D Gwennaelle Feinard; Quadiri Timour. 5-Fluorouracil-Induced Tako-Tsubo-like Syndrome. **2011**, 31 (2), 226–226. <https://doi.org/10.1592/phco.31.2.226>.
 30. Grunwald, M. R.; Howie, L.; Diaz, L. A. Takotsubo Cardiomyopathy and Fluorouracil: Case Report and Review of the Literature. *Journal of Clinical Oncology: official journal of the American Society of Clinical Oncology* **2012**, 30 (2), e11–4. <https://doi.org/10.1200/JCO.2011.38.5278>.
 31. Knott, K.; Starling, N.; Rasheed, S.; Foran, J.; Cafferkey, C.; Rosen, S.; Nicholson, A.; Baksi, J.; Lyon, A. A Case of Takotsubo Syndrome Following 5-Fluorouracil Chemotherapy. *International Journal of Cardiology* **2014**, 177 (2), e65–e67. <https://doi.org/10.1016/j.ijcard.2014.09.154>.
 32. Tsibiribi, P.; Bui-Xuan, C.; Bui-Xuan, B.; Lombard-Bohas, C.; Duperret, S.; Belkhiria, M.; Tabib, A.; Maujean, G.; Descotes,

- J.; Timour, Q. Cardiac Lesions Induced by 5-Fluorouracil in the Rabbit. *Human & Experimental Toxicology* **2006**, *25* (6), 305–309. <https://doi.org/10.1191/0960327106ht628oa>.
33. Becker, K.; Erckenbrecht, J. F.; Aussinger, D.; Frieling, T. Cardiotoxicity of the Antiproliferative Compound Fluorouracil. *Drugs* **1999**, *57* (4), 475–484. <https://doi.org/10.2165/00003495-199957040-00003>.
34. Gianni, M.; Dentali, F.; Grandi, A. M.; Sumner, G.; Hiralal, R.; Lonn, E. Apical Ballooning Syndrome or Takotsubo Cardiomyopathy: A Systematic Review. *European Heart Journal* **2006**, *27* (13), 1523–1529. <https://doi.org/10.1093/eurheartj/ehl032>.
35. Bhardwaj, P. V.; Chaubey, V. K.; Islam, A. M. Capecitabine-Induced Takotsubo Cardiomyopathy: A Case Report. *The Permanente Journal* **2019**, *23*. <https://doi.org/10.7812/tpp/18.245>.
36. Qasem, A.; Bin Abdulhak, A. A.; Aly, A.; Moormeier, J. Capecitabine-Induced Takotsubo Cardiomyopathy: A Case Report and Literature Review. *American Journal of Therapeutics* **2016**, *23* (5), e1188–e1192. <https://doi.org/10.1097/mjt.0000000000000134>.
37. Abdelmaseih, R.; Blanco, A.; Abdelmasih, R.; Desai, K.; Pothan, J.; Patel, J.; Balaraman, R. Cardio-Oncology: Capecitabine Can Sometimes Be Heartbreaking Capecitabine-Induced Takotsubo Cardiomyopathy Case Report and Literature Review. *Current problems in cardiology* **2021**, *46* (8), 100854. <https://doi.org/10.1016/j.cpcardiol.2021.100854>.
38. Plana, J. C.; Galderisi, M.; Barac, A.; Ewer, M. S.; Ky, B.; Scherrer-Crosbie, M.; Ganame, J.; Sebag, I. A.; Agler, D. A.; Badano, L. P.; Banchs, J.; Cardinale, D.; Carver, J.; Cerqueira, M.; DeCara, J. M.; Edvardsen, T.; Flamm, S. D.; Force, T.; Griffin, B. P.; Jerusalem, G. Expert Consensus for Multimodality Imaging Evaluation of Adult Patients during and after Cancer Therapy: A Report from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *Journal of the American Society of Echocardiography* **2014**, *27* (9), 911–939. <https://doi.org/10.1016/j.echo.2014.07.012>.
39. Touyz, R. M.; Herrmann, J. Cardiotoxicity with Vascular Endothelial Growth Factor Inhibitor Therapy. *npj Precision Oncology* **2018**, *2* (1). <https://doi.org/10.1038/s41698-018-0056-z>.
40. Mihalcea, D.; Memis, H.; Mihaila, S.; Vinereanu, D. Cardiovascular Toxicity Induced by Vascular Endothelial Growth Factor Inhibitors. *Life* **2023**, *13* (2), 366. <https://doi.org/10.3390/life13020366>.
41. Franco, T. Takotsubo Cardiomyopathy in Two Men Receiving Bevacizumab for Metastatic Cancer. *Therapeutics and Clinical Risk Management* **2008**, *Volume 4*, 1367–1370. <https://doi.org/10.2147/tcrm.s3960>.
42. Stewart, T.; Pavlakakis, N.; Ward, M. Cardiotoxicity with 5-Fluorouracil and Capecitabine: More than Just Vasospastic Angina. *Internal Medicine Journal* **2010**, *40* (4), 303–307. <https://doi.org/10.1111/j.1445-5994.2009.02144.x>.
43. Saif, M. W.; Shah, M. M.; Shah, A. R. Fluoropyrimidine-Associated Cardiotoxicity: Revisited. *Expert Opinion on Drug Safety* **2009**, *8* (2), 191–202. <https://doi.org/10.1517/14740330902733961>.
44. Molteni, L. P.; Rampinelli, I.; Cergnul, M.; Scaglietti, U.; Paino, A. M.; Noonan, D. M.; Bucci, E. O.; Gottardi, O.; Albini, A. Capecitabine in Breast Cancer: The Issue of Cardiotoxicity during Fluoropyrimidine Treatment. *The Breast Journal* **2010**, *16*, S45–S48. <https://doi.org/10.1111/j.1524-4741.2010.01004.x>.
45. Y-Hassan, S.; Tornvall, P.; Törnerud, M.; Henareh, L. Capecitabine Caused Cardiogenic Shock through Induction of Global Takotsubo Syndrome. *Cardiovascular revascularization medicine: including molecular interventions* **2013**, *14* (1), 57–61. <https://doi.org/10.1016/j.carrev.2012.10.001>.
46. Endo, A.; Yoshida, Y.; Nakashima, R.; Takahashi, N.; Tanabe, K. Capecitabine Induces Both Cardiomyopathy and Multifocal Cerebral Leukoencephalopathy. *International Heart Journal* **2013**, *54* (6), 417–420. <https://doi.org/10.1536/ihj.54.417>.
47. Abdelrahman, M.; McCarthy, M. T.; Haliana Yusuf, Osman, N. Successful Capecitabine Rechallenge Following 5-Fluorouracil-Induced Takotsubo Syndrome. *Oxford Medical Case Reports* **2016**, *2016* (3), 47–50. <https://doi.org/10.1093/omcr/omv072>.
48. Khanji, M. Y.; Nolan, S.; Gwynne, S.; Pudney, D.; Ionescu, A. Tako-Tsubo Syndrome after Trastuzumab — an Unusual Complication of Chemotherapy for Breast Cancer. *Clinical Oncology* **2013**, *25* (5), 329–329. <https://doi.org/10.1016/j.clon.2012.12.007>.
49. Burgy, M.; Brossat, H.; Barthelemy, P.; Imperiale, A.; Trinh, A.; Hazam, C. A.; Bergerat, J.-P.; Mathelin, C. First Report of Trastuzumab Treatment after Postoperative Takotsubo Cardiomyopathy. *Anticancer research* **2014**, *34* (7), 3579–3582.
50. Lees, C.; Yazdan-Ashoori, P.; Jerzak, K. J.; Gandhi, S. Takotsubo Cardiomyopathy during Anti-HER2 Therapy for Metastatic Breast Cancer. *The Oncologist* **2018**, *24* (2), e80–e82. <https://doi.org/10.1634/theoncologist.2018-0285>.
51. Azzurra Irelli; Ceriello, L.; Patruno, L. V.; Tessitore, A.; Edoardo Allesse; Cannita, K.; Fabiani, D. Takotsubo Syndrome during Pertuzumab and Trastuzumab Therapy for HER2-Positive Metastatic Breast Cancer. *Biomedicine* **2024**, *12* (1), 179–179. <https://doi.org/10.3390/biomedicine12010179>.
52. Patel, V.; Tetyana Okan; Yee, J. L.; Patel, J. Trastuzumab-Induced Takotsubo Cardiomyopathy Leading to LV Thrombosis and Thromboembolic Stroke: A Case Report. *European Heart Journal - Case Reports* **2025**, *9* (2). <https://doi.org/10.1093/ehjcr/ytaf002>.
53. Numico, G.; Sicuro, M.; Silvestris, N.; Mozzicafreddo, A.; Trogu, A.; Malossi, A.; Cristofano, A.; Thiebat, B. Takotsubo Syndrome in a Patient Treated with Sunitinib for Renal Cancer. *Journal of Clinical Oncology: official journal of the American Society of Clinical Oncology* **2012**, *30* (24), e218–20. <https://doi.org/10.1200/JCO.2012.42.4911>.
54. Ovadia, D.; Esquenazi, Y.; Bucay, M.; Bachier, C. R. Association between Takotsubo Cardiomyopathy and Axitinib: Case Report and Review of the Literature. *Journal of Clinical Oncology: official journal of the American Society of Clinical Oncology* **2015**, *33* (1), e1–3. <https://doi.org/10.1200/JCO.2013.48.7280>.
55. Young Kwang Chae; Chiec, L.; Adney, S. K.; Waitzman, J. S.; Costa, R.; Carneiro, B. A.; Matsangou, M.; Agulnik, M.; Kopp, P.; Giles, F. Posterior Reversible Encephalopathy Syndrome and Takotsubo Cardiomyopathy Associated with Lenvatinib Therapy for Thyroid Cancer: A Case Report and Review. *Oncotarget* **2018**, *9* (46), 28281–28289. <https://doi.org/10.18632/oncotarget.25606>.
56. Fashafsha, Z.; Za, F.; Dy, S.; Df, M.; Es, T.; Aa, B.; Es, P.; Mg, P. *Annals of Clinical Case Reports Lenvatinib -Induced Takotsubo Syndrome in Patient with Thyroid Cancer OPEN ACCESS*; **2022**. <https://www.anncaserep.com/open-access/lenvatinib-ndash-induced-takotsubo-syndrome-in-patient-with-thyroid-cancer-9313.pdf> (accessed 2025-11-16).
57. Ramos, G. E.; Caglevic, C.; Bulnes, J. F.; Panay, S. E.; Zapata, M. I.; Daniele, A. J.; Rodríguez, M. E. Takotsubo Cardiomyopathy Afatinib-Related in a Non-Small Cell Lung Cancer Patient: Case Report. *Frontiers in cardiovascular medicine* **2022**, *9*, 1060813. <https://doi.org/10.3389/fcvm.2022.1060813>.
58. Giza, D. E.; Moudgil, R.; Lopez-Mattei, J.; Kim, P.; Iliescu, C. Association between Ibrutinib and Mid-Cavitary Takotsubo Cardiomyopathy: A Case Report and a Review of Chemotherapy-Induced Takotsubo's Cardiomyopathy. *European Heart Journal - Case Reports* **2017**, *1* (2). <https://doi.org/10.1093/ehjcr/ytx006>.

59. Tanaka, S.; Yagi, R.; Yoshimura, N.; Asakawa, R.; Ueno, K. Two Cases of Takotsubo Cardiomyopathy during Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitor Treatment for Rare Non-Small Cell Lung Cancer with Epidermal Growth Factor Receptor Mutation EGFR遺伝子変異を有する稀な非小細胞肺癌に対してEGFR-TKI投与中にたこつぼ型心筋症が生じた2症例. *Haigan* **2024**, *64* (3), 200–205. <https://doi.org/10.2482/haigan.64.200>.
60. Trontzas, I. P.; Vathiotis, I. A.; Kyriakoulis, K. G.; Sofianidi, A.; Spyropoulou, Z.; Charpidou, A.; Kotteas, E. A.; Syrigos, K. N.; ImmunoTTS Collaborative Group. Takotsubo Cardiomyopathy in Cancer Patients Treated with Immune Checkpoint Inhibitors: A Systematic Review and Meta-Summary of Included Cases. *Cancers* **2023**, *15* (9), 2637. <https://doi.org/10.3390/cancers15092637>.
61. Zito, C.; Manganaro, R.; Ciappina, G.; Spagnolo, C. C.; Racanelli, V.; Santarpia, M.; Silvestris, N.; Carerj, S. Cardiotoxicity Induced by Immune Checkpoint Inhibitors: What a Cardio-Oncology Team Should Know and Do. *Cancers* **2022**, *14* (21), 5403. <https://doi.org/10.3390/cancers14215403>.
62. Geisler, B. P.; Raad, R. A.; Esaian, D.; Sharon, E.; Schwartz, D. A. Apical Ballooning and Cardiomyopathy in a Melanoma Patient Treated with Ipilimumab: A Case of Takotsubo-like Syndrome. *The Journal for ImmunoTherapy of Cancer* **2015**, *3* (1). <https://doi.org/10.1186/s40425-015-0048-2>.
63. Tan, N. Y. L.; Anavekar, N. S.; Wiley, B. M. Concomitant Myopericarditis and Takotsubo Syndrome Following Immune Checkpoint Inhibitor Therapy. *BMJ case reports* **2020**, *13* (9), e235265. <https://doi.org/10.1136/bcr-2020-235265>.
64. Yamada, K.; Mizuki Ida-Ichikawa; Fujimoto, N.; Ishida, M.; Kaoru Dohi. Takotsubo Syndrome in a Cancer Patient Treated with a Combination of Anti-Cancer Drugs Including Immune Checkpoint Inhibitors: A Case Report. *European Heart Journal - Case Reports* **2024**, *8* (8). <https://doi.org/10.1093/ehjcr/ytac355>.
65. Serzan, M.; Rapisuwon, S.; Krishnan, J.; Chang, I. C.; Barac, A. Takotsubo Cardiomyopathy Associated with Checkpoint Inhibitor Therapy: Endomyocardial Biopsy Provides Pathological Insights to Dual Diseases. *JACC: CardioOncology* **2021**, *3* (2), 330–334. <https://doi.org/10.1016/j.jaccao.2021.02.005>.
66. Malley, T.; Watson, E. A Case of Takotsubo Cardiomyopathy after Chemotherapy. *Oxford Medical Case Reports* **2016**, *2016* (4), 55–58. <https://doi.org/10.1093/omcr/omw014>.
67. Coen, M.; Rigamonti, F.; Roth, A.; Koessler, T. Chemotherapy-Induced Takotsubo Cardiomyopathy, a Case Report and Review of the Literature. *BMC Cancer* **2017**, *17* (1). <https://doi.org/10.1186/s12885-017-3384-4>.
68. Moldovan, B.; Adam, I.-I.; Radu-Mihai Pisciă; Vlad Untaru; Stolica, D.; Șpac, A.; Modrigan, I.; Ursu, M.; Jupoiu, L.; Frâncu, A.; Pescaru, F.; Amir Hubeishie; Zolog, A.; Vecerzan, L. Takotsubo Cardiomyopathy after Cytoreductive Surgery and Hyperthermic Intraperitoneal Chemotherapy for a Recurrent Colon Cancer: A Life-Threatening Complication. *Diagnostics* **2024**, *14* (21), 2402–2402. <https://doi.org/10.3390/diagnostics14212402>.
69. Özbay, B.; Gürses, E.; Kemal, H. S.; Şimşek, E.; Kültürsay, H. Takotsubo Syndrome Early after Treatment due to Non Cardio-toxic Chemotherapy Agents. *Türk Kardiyoloji Dernegi arsivi : Turk Kardiyoloji Derneginin yayin organidir* **2021**, *49* (1), 72–75. <https://doi.org/10.5543/tkda.2020.25590>.
70. Pristipino, C. *Psychosomatic Medicine in Oncologic and Cardiac Disease Study (PSYCHONIC)*. Clinicaltrials.gov. <https://www.clinicaltrials.gov/study/NCT03437642> (accessed 2025-09-16).
71. Pelliccia, F.; Kaski, J. C.; Crea, F.; Camici, P. G. Pathophysiology of Takotsubo Syndrome. *Circulation* **2017**, *135* (24), 2426–2441. <https://doi.org/10.1161/circulationaha.116.027121>.
72. Wilkhoo, H. S.; Visvanathan, H.; Chatterjee, S.; Hussain, S.; Gundaniya, P.; Francis, V.; Singh, B.; Pillai, A. A. Current Clinical Perspectives on Takotsubo Syndrome: Comprehensive Analysis of Diagnosis, Management, and Pathophysiology. *International Journal of the Cardiovascular Academy* **2024**, *10*. <https://doi.org/10.4274/ijca.2024.63825>.
73. Arunachalam, K.; Gnanaguruparan, S.; Paulowski, J. Takotsubo Cardiomyopathy and LV Outflow Tract Obstruction after Initiation of Novel Oral Chemotherapy. *Rhode Island medical journal* (2013) **2020**, *103* (10), 40–43.
74. *The TAKO-TSUBO and Cancer Registry (TTAC)*. Clinicaltrials.gov. <https://www.clinicaltrials.gov/study/NCT01944826?cond=%22Broken%20Heart%20Syndrome%22&viewType=Table&rank=8> (accessed 2025-10-12).
75. Lyon, A. R.; Bossone, E.; Schneider, B.; Sechtem, U.; Citro, R.; Underwood, S. R.; Sheppard, M. N.; Figtree, G. A.; Parodi, G.; Akashi, Y. J.; Ruschitzka, F.; Filippatos, G.; Mebazaa, A.; Omerovic, E. Current State of Knowledge on Takotsubo Syndrome: A Position Statement from the Taskforce on Takotsubo Syndrome of the Heart Failure Association of the European Society of Cardiology. *European journal of heart failure* **2016**, *18* (1), 8–27. <https://doi.org/10.1002/ejhf.424>.
76. *Characterization of Primary and Secondary Stress Related Takotsubo (MAESTRO)*. Clinicaltrials.gov. <https://www.clinicaltrials.gov/study/NCT05793008?id=NCT05793008&rank=1> (accessed 2025-08-16).
77. *Psychosocial Support for Patients With Takotsubo Syndrome*. Clinicaltrials.gov. <https://www.clinicaltrials.gov/search?id=NCT05977049> (accessed 2025-10-02).
78. *EACVI Study on Multimodality Cardiovascular Imaging of Inflammatory Cardiovascular Diseases (EACVI-INFLAME)*. Clinicaltrials.gov. <https://www.clinicaltrials.gov/study/NCT07077304?id=NCT07077304&rank=1> (accessed 2025-09-16).
79. Adhikari, A.; Asdaq, S. M. B.; Al Hawaj, M. A.; Chakraborty, M.; Thapa, G.; Bhuyan, N. R.; Imran, M.; Alshammari, M. K.; Alshehri, M. M.; Harshan, A. A.; Alanazi, A.; Alhazmi, B. D.; Sreeharsha, N. Anticancer Drug-Induced Cardiotoxicity: Insights and Pharmacogenetics. *Pharmaceuticals (Basel, Switzerland)* **2021**, *14* (10), 970. <https://doi.org/10.3390/ph14100970>.
80. Zalewska-Adamiec, M.; Klonowska, P.; Małyszko, J.; Kuźma, Ł.; Bachorzewska-Gajewska, H.; Dobrzycki, S. Primary Takotsubo Syndrome as a Complication of Bladder Cancer Treatment in a 62-Year-Old Woman. *The American journal of case reports* **2021**, *22*, e930090. <https://doi.org/10.12659/AJCR.930090>.
81. Matsumoto, T.; Oda, T.; Yoshida, Y.; Kimura, S.; Hitomi Himi; Takao Tsuduki; Takagi, S.; Masahiro Takatani; Morishita, H. Takotsubo Cardiomyopathy Caused by Infusion Reaction to Trastuzumab. *International Cancer Conference Journal* **2019**, *9* (1), 28–31. <https://doi.org/10.1007/s13691-019-00396-0>.
82. Monti, M.; Cortesi, P.; Vespignani, R.; Ilaria Bronico; Gallio, C.; Flospergher, M.; Matteucci, L.; Giovanni Luca Frassinetti. Takotsubo Syndrome (TTS) in Onco-Hematologic Patients: Retrospective Analysis and Focus on the Correlation or Not with Anticancer Drugs. Case Reports and Review of the Literature. *Frontiers in Oncology* **2022**, *12*. <https://doi.org/10.3389/fonc.2022.875391>.
83. Fiste, O.; C. Trika; Syrigos, N. K.; Kotteas, E. A. A Rare Case of Takotsubo Cardiomyopathy. *PubMed* **2024**, *28* (5), 2063–2067. https://doi.org/10.26355/eurrev_202403_35619.
84. Taamallah, K.; Hela Ghedira; Younes Arous; Abdeddayem Hagui; Wafa Fehri. *Does emotion before a chemotherapy course break a heart? Cardiogenic shock complicating Takotsubo cardiomyopathy*. ResearchGate. <https://doi.org/10.21203/rs.3.rs-142578/v1>.

85. Khan, N. A. J.; Pacioles, T.; Alsharedi, M. Atypical Takotsubo Cardiomyopathy Secondary to Combination of Chemo-Immunotherapy in a Patient with Non-Small Cell Lung Cancer. *Cureus* **2020**, *12* (7), e9429. <https://doi.org/10.7759/cureus.9429>.
86. Stępień, K.; Nowak, K.; Pasieka, P.; Warmuz, K.; Stępień, A.; Nessler, J.; Zalewski, J. Typical Variant of Takotsubo Cardiomyopathy in Oncological Patients. Two Case Reports and Review of the Literature. *Folia medica Cracoviensia* **2020**, *60* (1), 45–54. <https://doi.org/10.24425/fmc.2020.133485>.
87. Goel, S. Chemotherapy-Induced Takotsubo Cardiomyopathy. *World Journal of Clinical Cases* **2014**, *2* (10), 565. <https://doi.org/10.12998/wjcc.v2.i10.565>.
88. Singh, S. B.; Harle, I. A. Takotsubo Cardiomyopathy Secondary in Part to Cancer-Related Pain Crisis: A Case Report. *Journal of Pain and Symptom Management* **2014**, *48* (1), 137–142. <https://doi.org/10.1016/j.jpainsymman.2013.09.002>.
89. Alekseev, I. E.; Lysovolenko, N. L.; Gavriluk, E. V.; Khvostovoi, V. V.; Lazareva, N. V.; Bykanova, A. M.; Grebenkin, E. N.; Stanoevich, I. V. Takotsubo Cardiomyopathy in Locally Advanced Rectosigmoid Junction Cancer: A Case Report. *Siberian journal of oncology* **2023**, *22* (5), 161–172. <https://doi.org/10.21294/1814-4861-2023-22-5-161-172>.
90. Ho, J. S. Y.; Ming Ng, T. Y.; Cen, S.; Sia, C.-H.; Chan, P.-F.; Yeo, T.-C. Metastatic Ovarian Cancer Presenting as Takotsubo Cardiomyopathy. *Heart and Mind* **2021**, *5* (2), 61–63. https://doi.org/10.4103/hm.hm_29_21.

■ Author

Pranav Maddineni is a driven junior student at Farmington High School. He aspires to be a cardiothoracic surgeon and actively volunteers in community health. This work was supervised by Dr. Prahlad Parajuli, a tumor immunologist at Wayne State University. He directs translational cancer studies and advises several clinical projects.

The Role of Amyloid- β and Tau Proteins in Alzheimer's Disease

Pranavi Nandyala, Jobin Varkey

Dougherty Valley High School, 10550 Albion Rd, San Ramon, CA, 94582, USA; pranavifive@gmail.com

ABSTRACT: Alzheimer's disease is a complex neurodegenerative disorder associated with the buildup of amyloid- β plaques and tau neurofibrillary tangles, but the exact relationship between them is unclear. Understanding their interaction is important since therapies targeting only amyloid- β have not been entirely effective. This review synthesizes studies to clarify whether amyloid- β and tau act independently or together in Alzheimer's disease. Based on prior clinical trials of anti-A β antibodies, it has been demonstrated that plaque reduction has had inconsistent cognitive improvement. Early-phase tau-targeted therapies, on the other hand, are currently being developed, but not yet widely effective or researched. Mechanistic studies and combination therapies have suggested that A β and tau have a combined effect, which supports dual-targeted interventions. Imaging studies also show that amyloid and tau have an interconnected effect. These findings show that neurodegeneration acceleration in Alzheimer's is due to both proteins instead of a single pathology. Developing early detection techniques through biomarkers, combined with therapies focused on both A β and tau, may improve clinical outcomes in terms of prevention and treating symptoms. This topic was chosen because Alzheimer's disease is one of the leading causes of dementia worldwide, and understanding the combined roles of amyloid- β and tau may contribute to the development of more effective treatments.

KEYWORDS: Genetics and Molecular Biology of Disease, Neurobiology, Alzheimer's Disease, Amyloid- β , Tau.

■ Introduction

Alzheimer's disease (AD) is one of the most complex neurodegenerative disorders, which is why its exact underlying causes are still not fully understood.¹ After decades of research, two key proteins have been identified as central to its pathology. These proteins are amyloid- β and tau. The question, however, is how these abnormalities occur and interact to accelerate neurodegeneration. This lack of clarity has created a significant research gap within Alzheimer's studies. Due to the lack of understanding surrounding these two proteins, the development of effective diagnostic, clinical, and therapeutic strategies has been significantly limited.

Previous studies have provided insights into the roles of amyloid- β and tau in Alzheimer's disease, but their findings often focus on different aspects. Research on amyloid- β has emphasized its early accumulation in the brain and potential role in initiating the neurodegenerative process, such as studies involving amyloid PET imaging that detected plaque build-up years before symptoms were detected.² On the other hand, studies on tau suggest that its presence is more closely linked with cognitive decline than amyloid- β .³ For instance, autopsy and tau PET imaging studies show that the spread of neurofibrillary tangles strongly correlates with memory impairment and disease severity.³ Overall, some studies support a sequential interaction between the proteins, while others indicate that their pathways are independent. So, despite these advances in research, the relationship between the two proteins remains unclear. Further, clinical trials targeting amyloid- β alone have had limited cognitive improvement, which shows the need for a perspective that considers the combined effects of these proteins.⁴

Clarifying the relationship between amyloid- β and tau proteins is needed to improve our understanding of AD progression. Current evidence shows that focusing on one protein alone has not resulted in effective therapies, further showing the importance of considering both of their effects. This review addresses a critical gap by combining the roles of both proteins to offer a more comprehensive understanding that can inform future research, early detection strategies, and the development of treatments.

■ Methodology

The objective of this study is to examine the roles of amyloid- β and tau proteins in the development and progression of Alzheimer's disease. Specifically, the study aims to evaluate whether amyloid- β accumulation and tau hyperphosphorylation act sequentially or independently in Alzheimer's disease. This review also aims to understand how these findings affect current and future therapeutic and clinical strategies.

Given that this study is a literature review, the goal is to synthesize findings from peer-reviewed scientific research that investigate amyloid- β and tau pathologies in Alzheimer's disease. No experimental manipulation, data collection from human subjects, or in-field studies were conducted. The data for this study were gathered through searches of academic databases. Peer-reviewed journal articles, systematic reviews, and meta-analyses were included. Only studies published in English and studies containing data and models regarding amyloid- β and tau pathologies were selected.

The primary sources included original peer-reviewed research articles that contained experimental findings, imaging results, or analyses on A β and tau in human subjects or animal models. Secondary sources included review articles, systemat-

ic analyses, and meta-analyses that showed trends, evaluated different hypotheses, or analyzed results from multiple studies.

Because this study is entirely a review of literature, no human or animal subjects were involved, so no ethical approval was required. The study did not use any physical tools or laboratory materials. All the data was gathered from publicly accessible academic sources.

■ Discussion

Background: Amyloid and Tau:

The amyloid precursor protein (APP) is a protein found in many cells in the body, but it is especially common in neurons. In the healthy brain, APP helps with neuron growth, communication between brain cells, and repair after injury.⁵ In healthy cells, this protein can be broken down, or cleaved, by enzymes in two different ways. These two ways are different because one way is healthy, whereas the other can become harmful.²

In the healthy, also known as a non-amyloidogenic pathway, APP is cut by an enzyme called α -secretase. This produces a fragment called soluble APP- α (sAPP α). sAPP α helps protect neurons and also supports memory and learning.⁶ In the amyloidogenic pathway (unhealthy), however, APP is cut by two different enzymes. These are β -secretase (also known as BACE1) and γ -secretase. This process produces short protein pieces called amyloid-beta (A β) peptides, including A β_{40} and A β_{42} .² A β_{40} is the more common form, but A β_{42} is more likely to build up and aggregate, which makes it especially harmful.²

On the other hand, tau is a protein that is mainly found inside neurons, where it helps stabilize microtubules. Microtubules are long, tube-like structures that act as the “tracks” for transporting nutrients, organelles, and other molecules within the cell.⁷ These microtubules are important in neurons because they have long extensions called axons, which need efficient transport to connect different parts of the brain. Tau assists in keeping the microtubules straight and strong by attaching to them, which allows materials to move smoothly and neurons to function properly.⁸ Tau can also be modified by adding phosphate groups through the process of phosphorylation. In healthy brains, this is strictly regulated and helps control tau’s interaction with microtubules. So, when tau works correctly, it supports neuron stability, communication, and overall brain health.⁷

However, in Alzheimer’s disease, these normal processes become disrupted. APP is processed through the amyloidogenic pathway, which leads to an excessive accumulation of A β peptides that can aggregate into plaques. At the same time, abnormal hyperphosphorylation of tau causes it to detach from microtubules and form neurofibrillary tangles, which hinders neuronal transport and contributes to neurodegeneration. Together, these pathological changes interfere with communication between neurons, which contributes to cognitive decline.

The Amyloid Hypothesis:

The amyloid hypothesis supports that the accumulation of amyloid- β (A β) plaques outside neurons is the initiating factor in Alzheimer’s disease.⁹ Under normal conditions, small

amounts of A β are produced and efficiently cleared by the brain. However, dysregulation in this process leads to excessive accumulation. When A β peptides, especially A β_{42} , build up in the brain, they can form small clusters called oligomers, which are prone to aggregation. These oligomers are highly toxic and interfere with communication between neurons at synapses.¹⁰ Synapses are the connections that allow neurons to send signals to each other. Over time, the oligomers grow into larger, insoluble plaques that are even visible in postmortem brains of Alzheimer’s patients.¹⁰

The amyloid hypothesis suggests that this buildup of A β triggers downstream harmful events inside neurons. One major effect is the abnormal modification of tau proteins. Normally, tau helps stabilize microtubules inside neurons, but exposure to A β oligomers can activate enzymes like glycogen synthase kinase-3 β (GSK-3 β) that add excessive phosphate groups to tau, a process known as hyperphosphorylation.¹¹ Hyperphosphorylated tau detaches from microtubules, causing them to collapse and form neurofibrillary tangles. These tangles disrupt intracellular transport, reduce neuron survival, and contribute to the cognitive decline seen in Alzheimer’s disease.¹²

Additionally, A β plaques can cause oxidative stress and inflammation. Microglial cells, which are the brain’s immune cells, try to clear the excess A β . However, this response can backfire by producing inflammatory molecules that damage surrounding neurons.⁵ Synapses are weakened, neurons die, and the brain experiences the neurodegeneration that is commonly seen in Alzheimer’s disease. According to the amyloid hypothesis, this sequence from A β accumulation to tau pathology to neuron death is the cascade that ultimately causes the symptoms of Alzheimer’s, which include memory loss, confusion, and impaired reasoning.¹⁰

Studies supporting the amyloid hypothesis use two strong lines of evidence: genetic mutations in familial Alzheimer’s disease and imaging studies of amyloid- β (A β) in living brains. First, rare inherited mutations in genes such as APP, PSEN1, and PSEN2 have been associated with early-onset Alzheimer’s disease. These mutations increase the ratio of the A β_{42} form (which aggregates more easily) to A β_{40} , indicating that excessive and abnormal amyloid production can initiate disease processes. Secondly, advanced positron emission tomography (PET) imaging studies show that high A β accumulation can be detected years before cognitive symptoms appear in individuals.¹³ For example, a large multicenter PET study found that cognitively unimpaired participants who were both A β -PET positive and tau-PET positive had a much higher risk of progression, mild cognitive impairment (MCI), and dementia compared to those who were negative.¹³ Further, recent studies show that a high A β -PET signal predicts future cognitive and functional decline in clinically normal older adults.¹⁴ Overall, these findings reinforce the idea that A β accumulation is likely an early cause of the cascade from A β buildup to tau pathology to the death of neurons.

Despite decades of research focused on reducing amyloid- β (A β) plaques, the outcomes from many clinical trials have been low. This has raised questions about the validity of the amyloid cascade hypothesis in guiding Alzheimer’s disease therapies.

One major criticism is that while many anti-A β treatments successfully reduce plaque load or lower A β biomarkers, this reduction often does not correlate with meaningful cognitive or functional improvement for patients.^{4,15} For example, although plaque reduction has been achieved with monoclonal antibodies like aducanumab and lecanemab, the clinical benefit of slowing cognitive decline remains inconsistent.¹⁶ Another significant issue is regarding the timing of intervention and the disease stage of the patient. Many anti-amyloid trials were conducted in participants who already had mild to moderate dementia, a stage where amyloid has likely triggered the cascade. So, it is likely that the buildup of tau neurofibrillary tangles, synaptic loss, and neurodegeneration has already taken place. In these situations, removing A β may not significantly reverse damage that has already occurred.¹⁷ Accordingly, reviews have suggested that interventions in later stages of Alzheimer's are less effective and that earlier intervention in preclinical or prodromal stages may be required.¹⁸ These issues have led many researchers to conclude that focusing solely on A β is too narrow. Evidence from clinical trials and biomarker analyses has suggested that amyloid removal does not always prevent cognitive decline, raising questions about whether amyloid is the main cause of Alzheimer's or one of the many significant causes of the disease.^{17,18}

The Tau Pathology Hypothesis:

In Alzheimer's disease, tau becomes abnormally modified through the process of hyperphosphorylation. During this process, multiple phosphate groups are added to amino acids that make up tau, which include serine, threonine, and tyrosine. This modification is typically regulated by glycogen synthase kinase-3 β , cyclin-dependent kinase 5, and microtubule affinity-regulating kinase, all of which become dysregulated in Alzheimer's patients.¹² Hyperphosphorylation changes tau's normal conformation, which reduces its ability to bind to microtubules. As tau detaches from these microtubules, the structure of the microtubules weakens, which causes microtubule destabilization and hence the weakening of axon transport pathways. So, the cells' transport system for transporting organelles, vesicles, and other essential proteins along the axon is damaged. Mitochondria, synaptic vesicles, and neurotrophic factors fail to reach the synapses due to the impaired microtubules, which result in weakened neurotransmission and synaptic dysfunction.¹² These factors contribute directly to cognitive decline. The detached tau proteins self-aggregate to form soluble oligomers that are neurotoxic and eventually aggregate into insoluble neurofibrillary tangles (NFTs) within the soma and dendrites.⁷ The NFTs disrupt signaling pathways, cause oxidative stress, and trigger inflammatory responses from neighboring cells. Over time, the accumulation of NFTs leads to neuronal atrophy and cell death, which increases neurodegeneration across the brain's regions.⁷

The tau pathology hypothesis specifically claims that these tau tangles are more directly correlated with cognitive decline and symptom severity than amyloid- β plaques. Research has shown that the amount and location of tau tangles in the brain correspond closely with regions responsible for mem-

ory, reasoning, and other higher cognitive functions, such as the hippocampus, entorhinal cortex, and neocortex.^{3,19} For example, patients with high tau accumulation in the hippocampus often show severe memory loss, while individuals with significant amyloid plaque deposits may not yet show severe cognitive symptoms.¹⁹ This has led researchers to believe that tau pathology may be a better indicator of the progression and severity of Alzheimer's disease than amyloid alone.¹⁹

The tau hypothesis emphasizes that the presence and spread of the tangles closely align with the cognitive decline observed in patients, showing that tau is an important indicator for understanding and potentially treating Alzheimer's disease.^{3,20} Once tau proteins become abnormal and form tangles inside neurons, they don't just stay in one cell. Instead, evidence shows that tau can spread from one neuron to another.²¹ Abnormal tau acts as a template, causing normal tau proteins in neighboring neurons to misfold and aggregate. This propagation explains why tau pathology follows a predictable pattern in the brain. In Alzheimer's disease, tau tangles typically appear first in the entorhinal cortex, which is critical for memory formation. From there, tau spreads to the hippocampus and eventually to widespread regions of the neocortex, correlating with progressive memory loss, cognitive decline, and behavioral changes.¹⁹ Understanding how tau propagates is essential because it shows why targeting tau early, before widespread spread occurs, could be more effective in slowing or preventing cognitive decline. By slowing the transmission of misfolded tau between neurons, therapies could simultaneously treat symptoms and potentially slow the progression of Alzheimer's disease.²²

The role of tau in Alzheimer's disease has been strongly supported by multiple types of evidence, including neuroimaging studies, post-mortem analyses, and emerging tau-targeted therapies. Imaging studies have been essential in tracking tau accumulation in living patients. By using positron emission tomography (PET) scans with tau-specific tracers such as 18F-flortaucipir, researchers can visualize tau deposition in different brain regions.²³ These studies consistently show that tau accumulation correlates closely with cognitive decline, particularly in memory and executive function, whereas amyloid plaques alone often show a weaker correlation.²⁴ Longitudinal imaging studies also suggest that tau spreads in a predictable pattern, which is from the entorhinal cortex to the hippocampus and neocortex.^{3,19} Autopsy findings provide additional support for tau's role in Alzheimer's disease. The examination of post-mortem brains reveals that the amount of neurofibrillary tangles strongly correlates with the extent of the disease, even more than the density of amyloid plaques.²⁵ Studies also show that individuals with high amyloid- β levels but low tau often remain cognitively normal, further showing tau as a more pressing cause of the symptoms.²⁶

Amyloid vs. Tau - Are They Competing or Linked?:

One of the most prominent debates in Alzheimer's research is how amyloid- β (A β) and tau proteins interact in the brain. The amyloid cascade hypothesis states that A β accumulation occurs first and triggers downstream tau pathology, ultimately

leading to the death of neurons and, therefore, cognitive decline.^{9,10} According to this model, amyloid plaques act as an initiator that initiates a cascade of molecular events, including tau hyperphosphorylation, synaptic dysfunction, and neuroinflammation. Experimental studies in transgenic mouse models support this idea. The mice engineered to overexpress human APP develop A β plaques first, followed by tau tangles and neuronal loss.²⁷ These findings suggest that A β may act as a trigger for tau pathology.

However, some research challenges this sequential view. Observations in both human and animal models indicate that tau pathology can sometimes occur without significant amyloid accumulation, particularly in brain regions like the entorhinal cortex.²⁸ This has led to the concept of independent pathways, where amyloid and tau may contribute to neurodegeneration separately, and their combined presence accelerates disease progression. In this scenario, targeting one protein without addressing the other explains why many anti-amyloid therapies have had limited clinical improvement.^{10,13}

More recent studies propose a combined model, where A β and tau interact in a feedback loop. In this model, early amyloid deposition promotes tau misfolding and spread, which in turn intensifies amyloid toxicity, creating a constant cycle of neurodegeneration.²⁹ Misfolded tau can increase synaptic dysfunction and neuroinflammation, which then leads to further amyloid accumulation.³ This model is supported by longitudinal imaging and biomarker studies. These show that amyloid accumulation occurs before detectable tau pathology, but tau spread correlates more closely with clinical symptoms. Therefore, this reinforces the idea that both proteins are critical in Alzheimer's and that their roles are interdependent.^{3,30}

Understanding whether amyloid triggers tau or whether they act independently is crucial for developing effective therapies. If amyloid truly initiates tau pathology, then early anti-amyloid interventions could prevent downstream tau tangles. In contrast, if tau can propagate independently, then therapies must target both proteins or focus on tau directly to halt cognitive decline.³¹

Positron emission tomography (PET) imaging using A β tracers shows plaques in cognitively normal older adults, indicating that amyloid deposition can begin 15-20 years before noticeable memory decline.^{30,32} Similarly, cerebrospinal fluid (CSF) analysis shows decreased A β_{42} levels early in the pre-clinical stage, which correlates with the initial deposition of plaques in the brain.³³ In contrast, tau pathology appears closer to the onset of cognitive symptoms. Tau PET imaging and autopsy studies demonstrate that neurofibrillary tangles begin in the entorhinal cortex and hippocampus before spreading to the neocortex.^{24,34} These findings reinforce the idea that while amyloid may initiate a pathological cascade, tau accumulation is more closely tied to functional impairment. Additionally, there is an increasing amount of evidence supporting the overlap between the processes. Some studies show that tau deposition can begin in specific brain regions even when amyloid levels are low, suggesting that tau may also have independent triggers or be influenced by other factors such as aging, inflammation, or vascular changes.^{25,35} In summary, amyloid pathology gen-

erally precedes tau pathology, but tau's appearance and spread are more predictive of symptom onset and disease progression, overall proving that both amyloid and tau are closely linked to the progression of Alzheimer's Disease, rather than just a single protein pathology.

Therapeutic Implications and Clinical Trials:

Over the past decade, anti-amyloid therapies have been the primary focus of clinical trials aiming to slow Alzheimer's disease progression. Monoclonal antibodies such as aducanumab and lecanemab have been designed to selectively target amyloid- β (A β) plaques and assist in clearing them from the brain.^{30,36} Aducanumab was approved by the FDA in 2021 under accelerated approval, based on its ability to reduce amyloid plaques, although its clinical benefits on cognitive outcomes remain debated.^{30,36} Similarly, lecanemab has shown promising results in Phase 3 trials, demonstrating modest slowing of cognitive decline in early Alzheimer's patients.^{37,38} The results that these treatments have shown, along with donanemab, can be seen in Table 1.³⁹

Table 1: Clinical trials targeting amyloid- β consistently show effective plaque reduction; however, cognitive outcomes remain inconsistent. These findings suggest that amyloid removal alone may not be sufficient to significantly alter disease progression, especially in later stages of Alzheimer's disease.

Therapy	Trial Phase	Target/Mechanism	Outcome Summary	References
Aducanumab	3	Selective A β plaque removal via monoclonal antibody	Reduced amyloid burden on PET but had inconsistent cognitive benefits ^{30,36}	Sevigny et al., 2018; Knopman et al., 2021
Lecanemab	3	Targets soluble A β protofibrils	Slowed cognitive decline by 27% in early AD ^{37,38}	Swanson et al., 2021; van Dyck et al., 2023
Donanemab	3	Binds pyroglutamate, which is modified A β	Reduced plaques and slowed cognitive decline ³⁹	Mintun et al., 2021

The growing recognition of tau's importance has also led to the development of tau-focused therapies. Current approaches include monoclonal antibodies targeting pathological tau, which are small molecules that inhibit tau aggregation or phosphorylation.³¹ There are also strategies to stabilize microtubules disrupted by tau tangles.³¹ Early-phase clinical trials are testing compounds such as gosuranemab and tilavonemab, which aim to reduce tau propagation and improve cognitive outcomes in patients with mild Alzheimer's disease.³¹ Other approaches involve microtubule stabilizers, which aim to counteract the loss of tau function and maintain neuronal cytoskeleton integrity, thereby reducing synaptic dysfunction.⁴⁰ The results of these approaches are summarized in Table 2. While tau-targeted therapies are still largely in early-phase clinical trials, evidence suggests that they may provide cognitive benefits when administered during the early disease stage.³¹

Table 2: Tau-targeting therapies have shown limited clinical success, with most trials failing to demonstrate cognitive improvement. This suggests that timing and combination approaches may be more critical.

Therapy	Trial Phase	Mechanism	Outcome summary	References
Gosuranemab	2	Antibody against extracellular tau	Safe, but no significant cognitive benefit ³¹	Congdon & Sigurdsson, 2018
Tilavonemab	2	Prevents trans-synaptic tau spread	Did not meet efficacy endpoints ³¹	Congdon & Sigurdsson, 2018
TPI-287	1	Microtubule stabilizer	Limited efficacy, along with some safety issues ⁴⁰	Goldlust et al., 2024

Current therapies suggest that combining amyloid- and tau-targeted strategies may be more effective than focusing on either pathology alone. Clinical trials are beginning to explore sequential interventions, aiming to reduce amyloid deposition while also preventing tau-mediated neurodegeneration.⁴¹ This supports the idea that it is not amyloid or tau that individually accelerate neurodegeneration, but rather it is both of them. Studies targeting both amyloid-β and tau pathologies are summarized in the tables below, with Table 3 focused on therapeutic strategies and Table 4 focused on mechanistic evidence.

Table 3: The method of combining approaches that target amyloid-β and tau produces stronger effects on neurodegeneration than single-pathway approaches. These strategies support the view that Alzheimer’s disease requires multi-targeted approaches to slow progression.

Therapeutic Focus	Intervention Strategy	Outcome summary	References
Combination therapy	Anti-Aβ antibody and anti-tau antibody in sequential administration	Demonstrated that combination therapy had a large effect on reducing biomarkers of neurodegeneration ⁴¹	Cummings <i>et al.</i> , 2022
Mechanistic insight	Combined suppression of Aβ and tau in preclinical models	Found interdependent toxicity, meaning dual targeting improved neuronal function ⁴²	Busche & Hyman, 2020
Clinical trial perspective	Early amyloid clearance, followed by tau-targeted therapy	Proposed that early Aβ removal delays tau accumulation and clinical decline ⁴³	van Dyck <i>et al.</i> , 2023

Table 4 shows mechanistic evidence that demonstrates how amyloid-β and tau interact in Alzheimer’s. The studies performed using transgenic mouse models show that amyloid-β accumulation pushes tau to spread and therefore increases neuronal hyperactivity. So, when both proteins are inhibited, neuronal hyperactivity will be decreased. Cell culture and animal experiments have shown that Aβ-induced kinase activation influences tau. This shows that therapeutics targeting Aβ first, followed by tau, can slow the progression of Alzheimer’s disease. In addition, human PET imaging studies show that amyloid plaques can be consistently seen before NFTs form, which supports that their relationship is interlinked.

Table 4: Experimental and imaging studies provide strong evidence that amyloid-β and tau pathologies are biologically interconnected, with amyloid often accelerating tau spread. These findings support therapeutic strategies that intervene early in amyloid accumulation to prevent tau-related neurodegeneration.

Experimental Model	Outcome Summary	Implication	References
Transgenic mouse model	Aβ accumulation accelerates the spread of tau	Dual inhibition reduces neuronal hyperactivity ⁴²	Busche & Hyman, 2020
Cell culture and animal model	Tau pathology depends on Aβ-induced kinase activation	Sequential therapy, first targeting Aβ, followed by tau, may delay AD ⁴⁷	He <i>et al.</i> , 2018
PET imaging study on humans	The presence of amyloid precedes tau accumulation and predicts it	Supports the idea that early amyloid clearance is necessary to prevent tau accumulation ⁴⁸	Palmqvist <i>et al.</i> , 2020

Gaps in Knowledge:

One of the current gaps in Alzheimer’s research is that a significant proportion of elderly individuals have substantial amyloid-β (Aβ) deposition in their brains, but do not show severe cognitive impairment.⁴³ This suggests that the presence of amyloid plaques alone is not sufficient to cause dementia. Several explanations have been proposed to explain this phenomenon. First, cognitive reserve may play a protective role. Individuals with higher education levels and more engagement appear to tolerate higher amyloid loads without developing clinical symptoms, potentially due to compensatory neural

mechanisms that preserve cognitive function.⁴⁴ Second, co-existing pathologies may regulate the impact of amyloid. For example, tau pathology, neuroinflammation, and vascular dysfunction can interact with Aβ to determine whether cognitive symptoms occur. In some people, amyloid accumulation occurs without significant tau aggregation, possibly explaining why cognitive function remains intact.⁴⁵ Finally, the solubility of amyloid may matter. Soluble oligomeric forms of Aβ are more synaptotoxic than the insoluble fibrillar plaques that accumulate in the brain, meaning that individuals with primarily fibrillar deposits may experience less neuronal dysfunction.⁴⁶ Together, these findings go against the idea that amyloid is a straightforward factor in Alzheimer’s disease and show the importance of studying amyloid in combination with other pathological markers, including tau, to fully understand disease progression.

Implications for Future Research:

Research over the past decades has established that both amyloid-β (Aβ) and tau proteins play crucial roles in Alzheimer’s disease, but their relationship is complex. Aβ accumulation is often one of the earliest detectable changes in the brain and can initiate a cascade of events that promote tau hyperphosphorylation and neurofibrillary tangle formation.^{10,32} However, tau pathology correlates more closely with cognitive decline and symptom severity than amyloid burden, suggesting that tau tangles may be the more direct cause of the symptoms.³

Together, these findings suggest that Alzheimer’s is not caused by a single protein pathology. Rather, Aβ and tau interact with each other. The combination of tau and amyloid reveals why therapeutic strategies targeting only one protein have often produced mixed results, and why a more comprehensive understanding is essential for designing effective therapeutics and interventions. Given the interdependent relationship between amyloid-β and tau, there is a growing consensus that interventions in the preclinical or early stages of Alzheimer’s disease are likely to be more effective than treatments administered after significant neurodegeneration has occurred. Early detection through biomarkers, including cerebrospinal fluid (CSF) Aβ and tau measurements, PET imaging, and emerging blood tests, can identify individuals at risk before cognitive symptoms appear.⁴⁷⁻⁴⁹ With the combination of early detection and a treatment for both protein pathologies, more effective treatments can be developed. By focusing more on prevention and early intervention, and also by recognizing the interdependence between amyloid and tau, researchers can aim to slow disease progression, offering the potential to improve Alzheimer’s in the overall at-risk populations.²³

Conclusion

In summary, both amyloid-β and tau proteins play key roles in Alzheimer’s disease, but their relationship is not yet fully understood. While amyloid accumulation may initiate pathological cascades, tau pathology correlates more closely with cognitive decline, suggesting that both contribute to disease progression. This complexity shows the importance of early

detection using biomarkers and imaging techniques to identify at-risk individuals before significant neurodegeneration occurs.

The overall findings indicate that effective treatment may require a combined therapeutic approach that targets both amyloid and tau, alongside addressing other contributing factors such as inflammation and vascular changes. Future research should further investigate how these other factors interact with protein pathologies. Future research should also explore new biomarkers and intervention strategies to improve early detection and prevention of disease progression. Additionally, future studies should prioritize multi-model approaches, such as combining PET imaging, blood-based biomarkers, and genetic analyses, to better understand how amyloid and tau change over time. Researchers should also investigate combination therapies administered during preclinical stages of Alzheimer's disease to find whether early intervention can prevent cognitive decline. By researching various aspects of Alzheimer's disease, researchers can develop more effective methods to slow progression and hence improve patient treatments.

■ Acknowledgments

I sincerely thank Dr. Jobin Varkey and Professor Virgel Torremocha for guiding me through the process of my review and for their consistent mentorship. I greatly appreciate their guidance and feedback.

■ References

- Alzheimer's Association. *Alzheimer's Disease Facts and Figures*. Alzheimer's Disease and Dementia. <https://www.alz.org/alzheimers-dementia/facts-figures>.
- Haass, C.; Selkoe, D. J. Soluble Protein Oligomers in Neurodegeneration: Lessons from the Alzheimer's Amyloid β -Peptide. *Nature Reviews Molecular Cell Biology* **2007**, *8* (2), 101–112. <https://doi.org/10.1038/nrm2101>.
- Brier, M. R.; Gordon, B.; Friedrichsen, K.; McCarthy, J.; Stern, A.; Christensen, J.; Owen, C.; Aldea, P.; Su, Y.; Hassenstab, J.; Cairns, N. J.; Holtzman, D. M.; Fagan, A. M.; Morris, J. C.; Benzinger, T. L. S.; Ances, B. M. Tau and A β Imaging, CSF Measures, and Cognition in Alzheimer's Disease. *Science Translational Medicine* **2016**, *8* (338), 338ra66–338ra66. <https://doi.org/10.1126/scitranslmed.aaf2362>.
- Danko Jeremic; Navarro-Lopez, J. D.; Jimenez-Diaz, L. Clinical Benefits and Risks of Anti-Amyloid Antibodies in Sporadic Alzheimer's Disease: Systematic Review and Network Meta-Analysis with Web Application (Preprint). *Journal of Medical Internet Research* **2024**. <https://doi.org/10.2196/68454>.
- Muller, U. C.; Zheng, H. Physiological Functions of APP Family Proteins. *Cold Spring Harbor Perspectives in Medicine* **2011**, *2* (2), a006288–a006288. <https://doi.org/10.1101/cshperspect.a006288>.
- Kuhn, P.-H.; Wang, H.; Dislich, B.; Colombo, A.; Zeitschel, U.; Ellwart, J. W.; Kremmer, E.; Roßner, S.; Lichtenthaler, S. F. ADAM10 Is the Physiologically Relevant, Constitutive α -Secretase of the Amyloid Precursor Protein in Primary Neurons. *The EMBO Journal* **2010**, *29* (17), 3020–3032. <https://doi.org/10.1038/emboj.2010.167>.
- Wang, Y.; Mandelkow, E. Tau in Physiology and Pathology. *Nature Reviews Neuroscience* **2015**, *17* (1), 22–35. <https://doi.org/10.1038/nrn.2015.1>.
- Iqbal, K.; Liu, F.; Gong, C.-X.; Grundke-Iqbal, I. Tau in Alzheimer Disease and Related Tauopathies. *Current Alzheimer Research* **2010**, *7* (8), 656–664. <https://doi.org/10.2174/156720510793611592>.
- Hardy, J.; Higgins, G. Alzheimer's Disease: The Amyloid Cascade Hypothesis. *Science* **1992**, *256* (5054), 184–185. <https://doi.org/10.1126/science.1566067>.
- Selkoe, D. J.; Hardy, J. The Amyloid Hypothesis of Alzheimer's Disease at 25 Years. *EMBO molecular medicine* **2016**, *8* (6), 595–608. <https://doi.org/10.15252/emmm.201606210>.
- Hoover, B. R.; Reed, M. N.; Su, J.; Penrod, R. D.; Kotilinek, L. A.; Grant, M. K.; Pitstick, R.; Carlson, G. A.; Lanier, L. M.; Yuan, L.-L.; Ashe, K. H.; Liao, D. Tau Mislocalization to Dendritic Spines Mediates Synaptic Dysfunction Independently of Neurodegeneration. *Neuron* **2010**, *68* (6), 1067–1081. <https://doi.org/10.1016/j.neuron.2010.11.030>.
- Iqbal, K.; Liu, F.; Gong, C.-X. Tau and Neurodegenerative Disease: The Story so Far. *Nature Reviews Neurology* **2015**, *12* (1), 15–27. <https://doi.org/10.1038/nrneurol.2015.225>.
- Ossenkoppele, R.; Pichet Binette, A.; Groot, C.; Smith, R.; Strandberg, O.; Palmqvist, S.; Stomrud, E.; Tideman, P.; Ohlsson, T.; Jögi, J.; Johnson, K.; Sperling, R.; Dore, V.; Masters, C. L.; Rowe, C.; Visser, D.; van Berckel, B. N. M.; van der Flier, W. M.; Baker, S.; Jagust, W. J. Amyloid and Tau PET-Positive Cognitively Unimpaired Individuals Are at High Risk for Future Cognitive Decline. *Nature Medicine* **2022**, *28* (11), 2381–2387. <https://doi.org/10.1038/s41591-022-02049-x>.
- Nevra Öksüz; Reza Ghouri; Bahar Taşdelen; Derya Uludüz; Aynur Özge. Mild Cognitive Impairment Progression and Alzheimer's Disease Risk: A Comprehensive Analysis of 3553 Cases over 203 Months. *Journal of Clinical Medicine* **2024**, *13* (2), 518–518. <https://doi.org/10.3390/jcm13020518>.
- Chhabra, A.; Solanki, S.; Saravanabawan, P.; Venkiteswaran, A.; Nimmathota, N.; Modi, N. M. A Systematic Review of the Efficacy and Safety of Anti-Amyloid Beta Monoclonal Antibodies in Treatment of Alzheimer's Disease. *Expert opinion on biological therapy* **2024**, *24* (11), 1261–1269. <https://doi.org/10.1080/14712598.2024.2416947>.
- Avgerinos, K. I.; Manolopoulos, A.; Ferrucci, L.; Kapogiannis, D. Critical Assessment of Anti-Amyloid- β Monoclonal Antibodies Effects in Alzheimer's Disease: A Systematic Review and Meta-Analysis Highlighting Target Engagement and Clinical Meaningfulness. *Scientific Reports* **2024**, *14* (1). <https://doi.org/10.1038/s41598-024-75204-8>.
- Jin, M.; Noble, J. M. What's in It for Me? Contextualizing the Potential Clinical Impacts of Lecanemab, Donanemab, and Other Anti- β -Amyloid Monoclonal Antibodies in Early Alzheimer's Disease. *eneuro* **2024**, *11* (9), ENEURO.0088-24.2024. <https://doi.org/10.1523/eneuro.0088-24.2024>.
- Doran, S. J.; Sawyer, R. P. Risk Factors in Developing Amyloid Related Imaging Abnormalities (ARIA) and Clinical Implications. *Frontiers in Neuroscience* **2024**, *18*. <https://doi.org/10.3389/fnins.2024.1326784>.
- Vogel, J. W.; Young, A. L.; Oxtoby, N. P.; Smith, R.; Ossenkoppele, R.; Strandberg, O. T.; La Joie, R.; Aksman, L. M.; Grothe, M. J.; Iturria-Medina, Y.; Pontecorvo, M. J.; Devous, M. D.; Rabinovici, G. D.; Alexander, D. C.; Lyoo, C. H.; Evans, A. C.; Hansson, O. Four Distinct Trajectories of Tau Deposition Identified in Alzheimer's Disease. *Nature Medicine* **2021**, *27* (5), 871–881. <https://doi.org/10.1038/s41591-021-01309-6>.
- Goedert, M.; Spillantini, M. G. Pathogenesis of the Tauopathies. *Journal of Molecular Neuroscience* **2011**, *45* (3), 425–431. <https://doi.org/10.1007/s12031-011-9593-4>.

21. Clavaguera, F.; Bolmont, T.; Crowther, R. A.; Abramowski, D.; Frank, S.; Probst, A.; Fraser, G.; Stalder, A. K.; Beibel, M.; Staufenbiel, M.; Jucker, M.; Goedert, M.; Tolnay, M. Transmission and Spreading of Tauopathy in Transgenic Mouse Brain. *Nature Cell Biology* **2009**, *11* (7), 909–913. <https://doi.org/10.1038/ncb1901>.
22. Holmes, B. B.; Furman, J. L.; Mahan, T. E.; Yamasaki, T. R.; Mirbaha, H.; Eades, W. C.; Belaygorod, L.; Cairns, N. J.; Holtzman, D. M.; Diamond, M. I. Proteopathic Tau Seeding Predicts Tauopathy in *Vivo*. *Proceedings of the National Academy of Sciences* **2014**, *111* (41), E4376–E4385. <https://doi.org/10.1073/pnas.1411649111>.
23. Sperling, R.; Mormino, E.; Johnson, K. The Evolution of Preclinical Alzheimer's Disease: Implications for Prevention Trials. *Neuron* **2014**, *84* (3), 608–622. <https://doi.org/10.1016/j.neuron.2014.10.038>.
24. Ossenkoppele, R.; Schonhaut, D. R.; Schöll, M.; Lockhart, S. N.; Ayakta, N.; Baker, S. L.; O'Neil, J. P.; Janabi, M.; Lazaris, A.; Cantwell, A.; Vogel, J.; Santos, M.; Miller, Z. A.; Bettcher, B. M.; Vossel, K. A.; Kramer, J. H.; Gorno-Tempini, M. L.; Miller, B. L.; Jagust, W. J.; Rabinovici, G. D. Tau PET Patterns Mirror Clinical and Neuroanatomical Variability in Alzheimer's Disease. *Brain* **2016**, *139* (5), 1551–1567. <https://doi.org/10.1093/brain/aww027>.
25. Nelson, P. T.; Braak, H.; Markesbery, W. R. Neuropathology and Cognitive Impairment in Alzheimer Disease: A Complex but Coherent Relationship. *Journal of neuropathology and experimental neurology* **2009**, *68* (1), 1–14. <https://doi.org/10.1097/NEN.0b013e3181919a48>.
26. Sperling, R. A.; Aisen, P. S.; Beckett, L. A.; Bennett, D. A.; Craft, S.; Fagan, A. M.; Iwatsubo, T.; Jack, C. R.; Kaye, J.; Montine, T. J.; Park, D. C.; Reiman, E. M.; Rowe, C. C.; Siemers, E.; Stern, Y.; Yaffe, K.; Carrillo, M. C.; Thies, B.; Morrison-Bogorad, M.; Wagster, M. V. Toward Defining the Preclinical Stages of Alzheimer's Disease: Recommendations from the National Institute on Aging-Alzheimer's Association Workgroups on Diagnostic Guidelines for Alzheimer's Disease. *Alzheimer's & Dementia* **2011**, *7* (3), 280–292. <https://doi.org/10.1016/j.jalz.2011.03.003>.
27. Lewandowski, C. T.; Maldonado Weng, J.; LaDu, M. J. Alzheimer's Disease Pathology in APOE Transgenic Mouse Models: The Who, What, When, Where, Why, and How. *Neurobiology of Disease* **2020**, *139*, 104811. <https://doi.org/10.1016/j.nbd.2020.104811>.
28. Braak, H.; Del Tredici, K. The Preclinical Phase of the Pathological Process Underlying Sporadic Alzheimer's Disease. *Brain* **2015**, *138* (10), 2814–2833. <https://doi.org/10.1093/brain/awv236>.
29. Zhang, H.; Wei, W.; Zhao, M.; Ma, L.; Jiang, X.; Pei, H.; Cao, Y.; Li, H. Interaction between A β and Tau in the Pathogenesis of Alzheimer's Disease. *International Journal of Biological Sciences* **2021**, *17* (9), 2181–2192. <https://doi.org/10.7150/ijbs.57078>.
30. Jack, C. R.; Knopman, D. S.; Jagust, W. J.; Petersen, R. C.; Weiner, M. W.; Aisen, P. S.; Shaw, L. M.; Vemuri, P.; Wiste, H. J.; Weigand, S. D.; Lesnick, T. G.; Pankratz, V. S.; Donohue, M. C.; Trojanowski, J. Q. Tracking Pathophysiological Processes in Alzheimer's Disease: An Updated Hypothetical Model of Dynamic Biomarkers. *The Lancet Neurology* **2013**, *12* (2), 207–216.
31. Congdon, E. E.; Sigurdsson, E. M. Tau-Targeting Therapies for Alzheimer Disease. *Nature Reviews Neurology* **2018**, *14* (7), 399–415. <https://doi.org/10.1038/s41582-018-0013-z>.
32. Villemagne, V. L.; Burnham, S.; Bourgeat, P.; Brown, B.; Ellis, K. A.; Salvado, O.; Szoek, C.; Macaulay, S. L.; Martins, R.; Maruff, P.; Ames, D.; Rowe, C. C.; Masters, C. L. Amyloid β Deposition, Neurodegeneration, and Cognitive Decline in Sporadic Alzheimer's Disease: A Prospective Cohort Study. *The Lancet Neurology* **2013**, *12* (4), 357–367. [https://doi.org/10.1016/s1474-4422\(13\)70044-9](https://doi.org/10.1016/s1474-4422(13)70044-9).
33. Hansson, O.; Zetterberg, H.; Buchhave, P.; Londes, E.; Blennow, K.; Minthon, L. Association between CSF Biomarkers and Incipient Alzheimer's Disease in Patients with Mild Cognitive Impairment: A Follow-up Study. *The Lancet Neurology* **2006**, *5* (3), 228–234. [https://doi.org/10.1016/s1474-4422\(06\)70355-6](https://doi.org/10.1016/s1474-4422(06)70355-6).
34. Braak, H.; Braak, E. Neuropathological Staging of Alzheimer-Related Changes. *Acta Neuropathologica* **1991**, *82* (4), 239–259. <https://doi.org/10.1007/bf00308809>.
35. Duyckaerts, C.; Delatour, B.; Potier, M.-C. Classification and Basic Pathology of Alzheimer Disease. *Acta Neuropathologica* **2009**, *118* (1), 5–36. <https://doi.org/10.1007/s00401-009-0532-1>.
36. Sevigny, J.; Chiao, P.; Bussière, T.; Weinreb, P. H.; Williams, L.; Maier, M.; Dunstan, R.; Salloway, S.; Chen, T.; Ling, Y.; O'Gorman, J.; Qian, F.; Arastu, M.; Li, M.; Chollate, S.; Brennan, M. S.; Quintero-Monzon, O.; Scannevin, R. H.; Arnold, H. M.; Engber, T. The Antibody Aducanumab Reduces A β Plaques in Alzheimer's Disease. *Nature* **2016**, *537* (7618), 50–56. <https://doi.org/10.1038/nature19323>.
37. Swanson, C. J.; Zhang, Y.; Dhadda, S.; Wang, J.; Kaplow, J.; Lai, R. Y. K.; Lannfelt, L.; Bradley, H.; Rabe, M.; Koyama, A.; Reyderman, L.; Berry, D. A.; Berry, S.; Gordon, R.; Kramer, L. D.; Cummings, J. L. A Randomized, Double-Blind, Phase 2b Proof-of-Concept Clinical Trial in Early Alzheimer's Disease with Lecanemab, an Anti-A β Protofibril Antibody. *Alzheimer's Research & Therapy* **2021**, *13* (1). <https://doi.org/10.1186/s13195-021-00813-8>.
38. van Dyck, C. H.; Swanson, C. J.; Aisen, P.; Bateman, R. J.; Chen, C.; Gee, M.; Kanekiyo, M.; Li, D.; Reyderman, L.; Cohen, S.; Froelich, L.; Katayama, S.; Sabbagh, M.; Vellas, B.; Watson, D.; Dhadda, S.; Izzar, M.; Kramer, L. D.; Iwatsubo, T. Lecanemab in Early Alzheimer's Disease. *New England Journal of Medicine* **2022**, *388* (1), 9–21. <https://doi.org/10.1056/nejmoa2212948>.
39. Mintun, M. A.; Lo, A. C.; Duggan Evans, C.; Wessels, A. M.; Ardayio, P. A.; Andersen, S. W.; Shcherbinin, S.; Sparks, J.; Sims, J. R.; Brys, M.; Apostolova, L. G.; Salloway, S. P.; Skovronsky, D. M. Donanemab in Early Alzheimer's Disease. *New England Journal of Medicine* **2021**, *384* (18). <https://doi.org/10.1056/nejmoa2100708>.
40. Goldlust, S. A.; L Burt Nabors; Hsu, S.; Nimish Mohile; J Paul Duic; Benkers, T.; Singer, S.; Rao, M.; Cappello, L.; Silberman, S. L.; Farmer, G. Phase 1 Trial of TPI 287, a Microtubule Stabilizing Agent, in Combination with Bevacizumab in Adults with Recurrent Glioblastoma. *Neuro-Oncology Advances* **2024**. <https://doi.org/10.1093/oaajnl/vdae009>.
41. Cummings, J.; Lee, G.; Nahed, P.; Kambar, M. E. Z. N.; Zhong, K.; Fonseca, J.; Taghva, K. Alzheimer's Disease Drug Development Pipeline: 2022. *Alzheimer's & Dementia: Translational Research & Clinical Interventions* **2022**, *8* (1). <https://doi.org/10.1002/trc2.12295>.
42. Busche, M. A.; Hyman, B. T. Synergy between Amyloid- β and Tau in Alzheimer's Disease. *Nature Neuroscience* **2020**, *23* (10), 1183–1193. <https://doi.org/10.1038/s41593-020-0687-6>.
43. Jack, C. R.; Bennett, D. A.; Blennow, K.; Carrillo, M. C.; Dunn, B.; Haeberlein, S. B.; Holtzman, D. M.; Jagust, W.; Jessen, F.; Karlawish, J.; Liu, E.; Molinuevo, J. L.; Montine, T.; Phelps, C.; Rankin, K. P.; Rowe, C. C.; Scheltens, P.; Siemers, E.; Snyder, H. M.; Sperling, R. NIA-AA Research Framework: Toward a Biological Definition of Alzheimer's Disease. *Alzheimer's & Dementia: the Journal of the Alzheimer's Association* **2018**, *14* (4), 535–562. <https://doi.org/10.1016/j.jalz.2018.02.018>.
44. Stern, Y. Cognitive Reserve in Ageing and Alzheimer's Disease. *The Lancet Neurology* **2012**, *11* (11), 1006–1012. [https://doi.org/10.1016/s1474-4422\(12\)70191-6](https://doi.org/10.1016/s1474-4422(12)70191-6).
45. Bennett, D. A.; Schneider, J. A.; Arvanitakis, Z.; Kelly, J. F.; Aggarwal, N. T.; Shah, R. C.; Wilson, R. S. Neuropathology of Older Persons without Cognitive Impairment from Two Commu-

- nity-Based Studies. *Neurology* **2006**, 66 (12), 1837–1844. <https://doi.org/10.1212/01.wnl.0000219668.47116.e6>.
46. Benilova, I.; Karran, E.; De Strooper, B. The Toxic A β Oligomer and Alzheimer's Disease: An Emperor in Need of Clothes. *Nature Neuroscience* **2012**, 15 (3), 349–357. <https://doi.org/10.1038/nn.3028>.
47. CSF Biomarkers of Alzheimer's Disease Concord with Amyloid- β PET and Predict Clinical Progression: A Study of Fully Automated Immunoassays in BioFINDER and ADNI Cohorts. *Alzheimer's & Dementia* **2018**, 14 (11), 1470–1481. <https://doi.org/10.1016/j.jalz.2018.01.010>.
48. Janelidze, S.; Mattsson, N.; Palmqvist, S.; Smith, R.; Beach, T. G.; Serrano, G. E.; Chai, X.; Proctor, N. K.; Eichenlaub, U.; Zetterberg, H.; Blennow, K.; Reiman, E. M.; Stomrud, E.; Dage, J. L.; Hansson, O. Plasma P-Tau181 in Alzheimer's Disease: Relationship to Other Biomarkers, Differential Diagnosis, Neuropathology and Longitudinal Progression to Alzheimer's Dementia. *Nature Medicine* **2020**, 26 (3), 379–386. <https://doi.org/10.1038/s41591-020-0755-1>.
49. Palmqvist, S.; Janelidze, S.; Quiroz, Y. T.; Zetterberg, H.; Lopera, F.; Stomrud, E.; Su, Y.; Chen, Y.; Serrano, G. E.; Leuzy, A.; Mattsson-Carlsson, N.; Strandberg, O.; Smith, R.; Villegas, A.; Sepulveda-Falla, D.; Chai, X.; Proctor, N. K.; Beach, T. G.; Blennow, K.; Dage, J. L. Discriminative Accuracy of Plasma Phospho-Tau217 for Alzheimer Disease vs Other Neurodegenerative Disorders. *JAMA* **2020**, 324 (8), 772–781. <https://doi.org/10.1001/jama.2020.12134>.

■ Author

Pranavi Nandiyala is a junior at Dougherty Valley High School interested in science and medicine. She is especially passionate about researching brain disorders and understanding their causes to contribute to advances in neuroscience and healthcare.

Assessing the Impact of Microplastics on DNA Damage in *C. elegans*

Renee Ma

New Trier High School, 385 Winnetka Ave, Winnetka, IL 60093, USA; reneema16@gmail.com

ABSTRACT: DNA damage is a key driver of genetic mutations and disease, including cancer. Microplastics, generated from the degradation of plastic waste, are now widely detected in water, food, and the human body. However, their potential to induce DNA damage remains largely unknown. Here, we examined the effect of polystyrene microplastics on causing DNA damage using *Caenorhabditis elegans* (*C. elegans*) as a model organism. Transgenic *cep-1::GFP* worms, which express green fluorescent protein (GFP) under the control of the *cep-1* promoter in response to DNA damage, were exposed to polystyrene beads of two sizes (0.2 μm or 1 μm) at two concentrations (1×10^7 particles/mL and 1×10^8 particles/mL) for 24 hours. Wild-type and untreated transgenic worms served as controls. GFP expression was quantified by fluorescence microscopy and ImageJ analysis. Microplastic exposure significantly elevated GFP expression in *cep-1::GFP* worms regardless of microplastic size or concentration, suggesting robust induction of DNA damage response, whereas neither control group showed a significant response. Altogether, these results demonstrate that polystyrene microplastics can induce DNA damage in *C. elegans* and suggest that similar genotoxic effects may occur in other organisms, including humans, highlighting a potential link between microplastic exposure and disease pathogenesis.

KEYWORDS: Biology, Cellular and Molecular Biology, DNA Damage, Microplastics, *C. elegans*.

■ Introduction

Double-stranded deoxyribonucleic acid (DNA) serves as the genetic blueprint that determines protein and cellular functions in most organisms.¹ Mutations in DNA can disrupt these functions and contribute to a wide range of diseases, including cancer, autism, and degenerative disorders.² While some genetic mutations are inherited, most disease-causing mutations are not passed down through familial lines.^{3,4} These mutations could result from DNA damage induced by environmental stressors such as chemical exposure, radiation, and pollutants.³⁻⁵ This study investigates the genotoxic potential of microplastics - widespread contaminants found in food and water, and the human body - and their possible contribution to the pathogenesis of cancer and other diseases.

Microplastics are defined as plastic particles less than 5 mm in diameter.^{6,7} They are generated from the degradation of plastic waste or intentionally produced for use in paints and personal care products, such as toothpaste, sunscreen, and cosmetics.⁸ Among the 50 or more chemically distinct types of plastics, polystyrene, polyethylene, and polypropylene are commonly used through mass production. Over the past century, global plastic production has increased dramatically and now exceeds 400 million tons per year.⁹ This has led to the accumulation of a large amount of microplastics in environments.¹⁰ Recent studies have detected microplastics in various ecosystems, including oceans, freshwater bodies, and the air.¹¹ Microplastics have also been found in food sources, drinking water, and even human breastmilk,¹² placentas,¹³ blood,¹⁴ lungs,¹⁵ and brains,¹⁶ raising concerns about their potential toxicity.¹⁶⁻¹⁸ This study examines the ability of polystyrene microplastics, one of the

most widely used and potentially toxic plastics,¹⁹ to induce DNA damage in *C. elegans*.

C. elegans is a small nematode, approximately 1mm in length, commonly found in soil where it feeds on bacteria (Figure 1). The species is non-pathogenic and non-parasitic, making it a safe organism for laboratory research. *C. elegans*' unique combination of biological, genetic, and experimental advantages has established it as a widely used model organism.^{20,21} The adult worm contains approximately 1,000 somatic cells, facilitating detailed studies of cellular processes. Its transparent body allows direct visualization of these processes by using microscopy. *C. elegans* completes its life cycle in about 3 days, with a total lifespan of only 2-3 weeks, enabling quick experimental turnover. The majority of *C. elegans* adults are hermaphrodites, possessing both male and female reproductive structures, but hermaphrodites cannot mate with each other. Meanwhile, as males can mate with hermaphrodites, a genetic cross can be easily performed in laboratories. The organism is easily maintained on agar plates using *Escherichia coli* OP50 as a food source, reducing both time and cost of experimentation.²⁰ The fully sequenced *C. elegans* genome consists of roughly 20,000 genes, many of which are conserved with human orthologs. In addition, genetic manipulation in *C. elegans* is straightforward: RNA interference (RNAi) can be performed simply by feeding worms bacteria expressing double-stranded RNA, and CRISPR-Cas9 and other genome-editing tools can be used to generate desired mutants or modify specific genes.²¹ Many fundamental cellular and molecular pathways are conserved between *C. elegans* and humans, including programmed cell death, aging, and stress responses, making it a powerful model for studying human biology and disease.^{22,23} The small size and

ease of handling make *C. elegans* suited for high-throughput screening of potential therapeutics and genetic modifiers. As an invertebrate, *C. elegans* also raises fewer ethical concerns than vertebrate model organisms. Therefore, *C. elegans* offers a combination of biological simplicity, ease of genetic analysis, conservation of essential pathways, and cost-effectiveness. These qualities have made it an important model system of research in genetics, toxicology, developmental biology, and aging.²²⁻²⁴

The molecular mechanisms underlying the DNA damage response are conserved in *C. elegans*. The *cep-1* gene, the *C. elegans* ortholog of the human tumor suppressor *p53*, plays a central role in stress response and DNA damage repair, and can be activated in response to DNA damage in *C. elegans*.²⁵ Transgenic *C. elegans* that express GFP under the control of *cep-1* offers a simple, sensitive, and quantitative readout of DNA damage in a living organism.²⁵ Polyethylene microplastics have been shown to cause reproductive cell toxicity in *C. elegans*.²⁶ However, the effect of the potentially more toxic plastic, polystyrene, on DNA damage remains unknown.¹⁹ This study leverages the transgenic *cep-1::GFP* *C. elegans*, which expresses GFP in response to DNA damage, to investigate whether polystyrene microplastic beads can induce DNA damage. Microplastics of two different sizes and two concentrations are tested. We choose 0.2 μm because it represents the smallest end of the microplastics spectrum, and 1 μm nears the upper limit of what may enter cells.^{7,27} The two concentrations (1×10^7 particles/mL and 1×10^8 particles/mL) resemble the concentration of microplastics found in environmental surface water.²⁸ Findings from this study will reveal the potential role of polystyrene microplastics in contributing to DNA damage and genotoxicity.

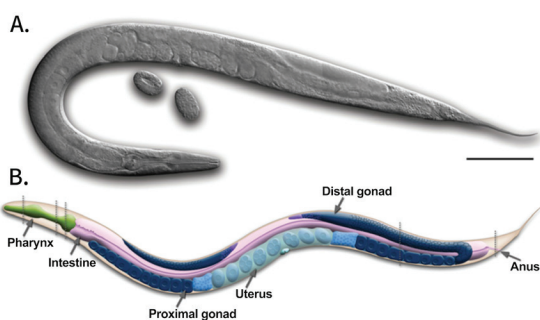


Figure 1: (A) Differential interference contrast (DIC) image of an adult hermaphrodite *C. elegans* (left lateral side) and two eggs. Scale bar 0.1 mm. (B) Schematic drawing of an adult hermaphrodite *C. elegans* anatomical structures, left lateral side. (Modified from WormAtlas.org)

■ Methods

Materials:

1. Polystyrene microplastics:

Polystyrene particles with diameters of 1 or 0.2 μm were obtained from Polysciences (#07310, #07304). Prior to use, microplastic beads were washed three times with M9 worm buffer by centrifuging (Eppendorf 5424) at 21,000 g for 5 minutes and discarding the supernatant. They were then suspended in M9 buffer as a 100x stock solution before being used to treat *C. elegans*.

2. *C. elegans* worms:

Wild-type N2 and transgenic *cep-1::GFP* (#TG12) worms were obtained from the *Caenorhabditis* Genetics Center (CGC) and maintained using standard protocols, feeding with *E. coli* OP50 on nematode growth medium plates in a 20°C incubator.⁵

3. M9 worm buffer and worm bleach:

M9 buffer preparation: (100mL)

1.5 g KH_2PO_4
3.0 g Na_2HPO_4
0.25 g NaCl
0.5 g NH_4Cl
Bring to 500 mL with H_2O .

Worm bleach preparation:

40 mL H_2O
40 mL bleach (5% hypochlorite)
20 mL 5N KOH
Cover bottle with aluminum foil and use within a month of making it.

4. Nematode growth medium plates and *E. coli* OP50:

Nematode growth medium plates are prepared by mixing 3 g NaCl, 17 g agar, 2.5 g peptone, and 975 mL H_2O in a flask. After autoclaving, add 1 mL of 1 M CaCl_2 , 1 mL of 5 mg/mL cholesterol in ethanol, 1 mL of 1 M MgSO_4 , and 25 mL of 1 M KPO_4 buffer. Dispense the mixture into sterile petri dishes. *E. coli* OP50 is cultured in standard LB medium at 37°C and stored at 4°C for several months.

Experimental Design and Methods:

1. Cultivation and synchronization of *C. elegans*:

C. elegans were cultivated on nematode growth medium plates seeded with *E. coli* OP50 as a food source. The health and density of worm populations were routinely monitored throughout the experiment. Wild-type and transgenic strains were synchronized using bleach treatment to minimize variation arising from developmental stage differences. For synchronization, worms were exposed to a worm bleach solution for three minutes. Adult worms were lysed by the bleach, while worm eggs, which are resistant to bleach, remained intact. After five washes, eggs were collected and transferred to fresh nematode growth medium plates, generating a new, synchronous population. This process ensured uniform developmental stages among the worms used for subsequent experiments (Figure 2).

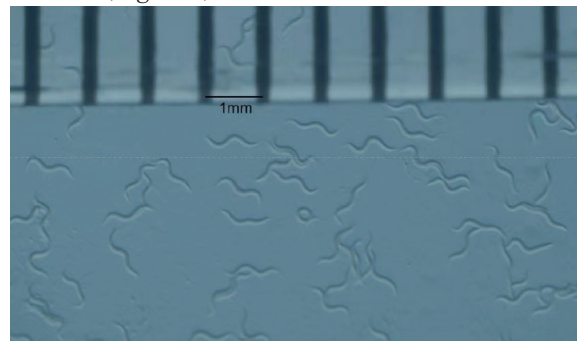


Figure 2: Synchronized wild-type N2 adult worms on a Nematode growth medium plate. Scale bar 1mm.

2. Microplastic Treatment:

Synchronized Day 1 adult worms were collected and transferred into 1.5 mL Eppendorf tubes containing M9 worm

buffer. Polystyrene microplastic beads of different sizes were added to the tubes to reach the desired concentration for each treatment group. Worms were incubated with polystyrene microplastic beads at 20°C for 24 hours.

The experiment consists of the following groups:

Group 1: Wild-type worms (no *GFP* transgene) not exposed to microplastics.

Group 2: Wild-type worms exposed to 0.2 μm (size) microplastics at 1×10^7 particles/mL.

Group 3: Transgenic *cep-1::GFP* worms not exposed to microplastics.

Group 4: Transgenic *cep-1::GFP* exposed to 0.2 μm (size) microplastics at 1×10^7 particles/mL.

Group 5: Transgenic *cep-1::GFP* exposed to 0.2 μm (size) microplastics at 1×10^8 particles/mL.

Group 6: Transgenic *cep-1::GFP* exposed to 1 μm (size) microplastics at 1×10^7 particles/mL.

3. Worm Mounting and Imaging:

Following experimental treatments, worms were washed three times with saline to remove residual microplastics and subsequently anesthetized using 10 mM (0.7%) sodium azide. The anesthetized worms were embedded in Aquamount mounting medium under coverslips on glass slides for microscopy. GFP expression was imaged using a Keyence BZ-X800 fluorescence microscope, maintaining consistent exposure settings for all samples to ensure reproducibility.

4. Utilize ImageJ to quantify fluorescence intensity:

Images were converted to 8-bit black and white mode using the free ImageJ software prior to analysis. To minimize background signals, a fixed threshold was applied consistently across all images. The rectangle tool was used to select the region containing each worm, and the total pixel intensity within the selected area was measured and recorded in arbitrary units as the GFP expression level. For each experimental group, fluorescence levels were quantified in at least fifteen worms or images, with analysis conducted in a blinded manner to reduce bias.

5. Statistical analysis:

Statistical analysis of GFP expression was conducted to assess the effects of polystyrene microplastic treatment. One-way ANOVA was performed to compare GFP expression levels among four groups: wild type, wild type with microplastics, *cep-1::GFP*, and *cep-1::GFP* with microplastics, with microplastic treatment serving as the independent variable. Statistical significance was determined based on group comparisons. Furthermore, Student's t-test was used to evaluate whether significant differences in GFP expression existed between *cep-1::GFP* worms treated with different concentrations of 0.2 μm polystyrene microplastic beads (Group 4 vs. Group 5) and between *cep-1::GFP* worms exposed to microplastic beads of varying sizes at the same concentration of 1×10^7 particles/mL (Group 4 vs. Group 6), considering microplastic concentration or size as the independent variables. The complete experimental procedure is illustrated in Figure 3.

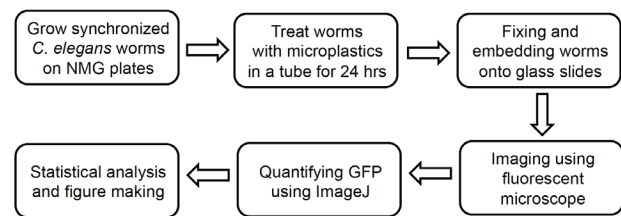


Figure 3: Flow chart for the experimental procedure.

Results and Discussion

The Effect of Microplastic Treatment:

Imaging and quantification of GFP expression revealed weak fluorescence in wild-type worms, both in untreated controls (Figure 4A, B) and in those exposed to 1×10^7 particles/mL 0.2 μm polystyrene microplastics (Figure 4C, D), showing background level fluorescence. Similarly, untreated *cep-1::GFP* transgenic worms displayed low GFP expression (Figure 4E, F), serving as a negative control. In contrast, a marked increase in GFP expression was observed in *cep-1::GFP* worms exposed to 1×10^7 particles/mL 0.2 μm microplastics ($P < 0.0001$) (Figure 4G, H; Figure 5). These results indicate that microplastic exposure activates *cep-1* and DNA damage response pathways, demonstrating that 0.2 μm polystyrene microplastics at 1×10^7 particles/mL induce DNA damage in *C. elegans*.

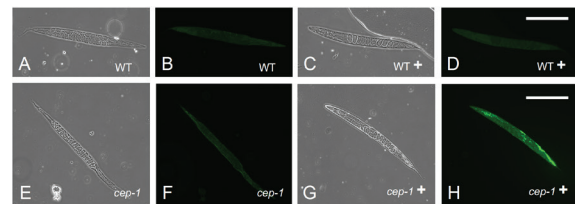


Figure 4: Microplastic treatment activates *cep-1* and DNA damage response pathways in *C. elegans*. (A, B) Bright field (A) and fluorescent (B) images of a wild type (WT) worm unexposed to microplastics, showing background level fluorescence. (C, D) WT worm exposed (+) to 0.2 μm polystyrene microplastics at 1×10^7 particles/mL, showing background level fluorescence. (E, F) Untreated *cep-1::GFP* transgenic worm showing baseline GFP expression. (G, H) *cep-1::GFP* transgenic worm exposed to 0.2 μm polystyrene microplastics at 1.66×10^{-14} M, showing strongly elevated GFP expression indicative of DNA damage response activation. Scale bar: 0.5 mm.

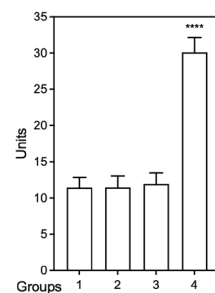


Figure 5: Quantification of GFP expression showing that polystyrene microplastic exposure induces robust GFP expression in *cep-1::GFP* transgenic worms (Group 4), but not in any other control groups. Group 1: WT worms unexposed to microplastics, 11.44 ± 1.40 Units, $n = 20$; Group 2: WT worms exposed to 0.2 μm polystyrene microplastics at 1×10^7 particles/mL, 11.45 ± 1.59 Units, $n = 20$; Group 3: *cep-1::GFP* worms unexposed to microplastics, 11.92 ± 1.54 Units, $n = 20$; Group 4: *cep-1::GFP* worms exposed to 0.2 μm microplastics at 1×10^7 particles/mL, 30.09 ± 2.05 Units, $n = 20$. Units: arbitrary units of GFP expression. Data are presented as mean $\pm$ SD. **** $P < 0.0001$, one-way ANOVA with Tukey's HSD *post hoc* test was performed on Groups 1-4. Data are from four independent biological/experimental replicates.

The Effects of Microplastic Concentration and Particle Size:

To evaluate the impact of microplastic concentration, *cep-1::GFP* transgenic worms were exposed to 0.2 μm microplastics at two concentrations: 1×10^7 particles/mL and 1×10^8 particles/mL. Both concentrations led to robust *cep-1* activation and GFP expression. Importantly, both concentrations produced a significant rise in GFP expression compared to the control. These findings indicate that the DNA damage response triggered by microplastic exposure operates in a threshold-dependent, binary manner; once this threshold is exceeded, increasing microplastic concentration does not further enhance the magnitude of the DNA damage response.

To examine the effect of microplastic size on DNA damage, *cep-1::GFP* transgenic worms were exposed to polystyrene microplastic beads of two sizes (0.2 μm or 1 μm ; both at 1×10^7 particles/mL; Group 4 vs. Group 6). Both sizes induced robust *cep-1* activation and GFP expression. However, worms exposed to smaller (0.2 μm) microplastic beads exhibited stronger GFP expression than those exposed to larger (1 μm) beads ($P = 0.02$). This is likely because smaller microplastic particles may be easier to enter *C. elegans* and its cells, or smaller particles have different surface chemistry or physics that make them more toxic.

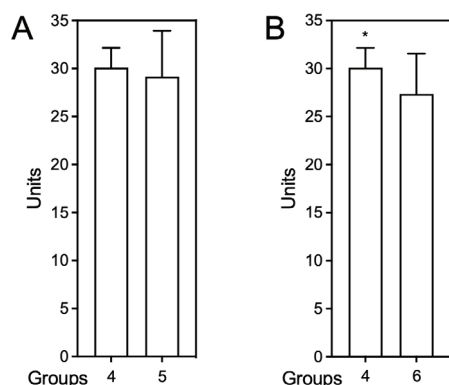


Figure 6: Both concentrations of microplastics induce robust *cep-1* activation and GFP expression, with smaller microplastics showing greater efficiency. **A.** Quantification of GFP expression showing no significant difference between two distinct concentrations of 0.2 μm polystyrene microplastics in inducing GFP expression in *cep-1::GFP* transgenic worms. Group 4: *cep-1::GFP* worms exposed to 0.2 μm polystyrene microplastic beads at 1×10^7 particles/mL, 30.09 ± 2.05 Units, $n = 20$; Group 5: *cep-1::GFP* worms exposed to 0.2 μm microplastics at 1×10^8 particles/mL, 29.17 ± 4.78 Units, $n = 15$. Units: arbitrary units of GFP expression quantified by Image J. Data are presented as mean $\pm$ SD. $P = 0.44$, Student's *t*-test. **B.** Quantification showing stronger GFP expression in *cep-1::GFP* worms exposed to 0.2 μm (size) microplastic beads than worms exposed to 1 μm microplastic beads, both at 1×10^7 particles/mL concentration. Group 4: *cep-1::GFP* worms exposed to 0.2 μm microplastics at 1×10^7 particles/mL, 30.09 ± 2.05 Units, $n = 20$; Group 6: *cep-1::GFP* worms exposed to 1 μm microplastics at 1×10^7 particles/mL, 27.36 ± 4.19 Units, $n = 15$. Data are presented as mean $\pm$ SD. $P = 0.02$, Student's *t*-test. Data are from four independent biological/experimental replicates.

Discussions:

This study examined polystyrene microplastics of two different sizes (0.2 μm or 1 μm) and at two concentrations (1×10^7 particles/mL and 1×10^8 particles/mL). These particle sizes were chosen because 0.2 μm is close to the smallest end of the microplastics spectrum, bordering on nanoplastics, and 1

μm is near the upper limit of what may enter cells.^{7,27} The two concentrations were tested because they resemble the concentration of microplastics found in environmental surface water.²⁸ Findings from this study reveal no significant difference between the two concentrations of polystyrene microplastics in their abilities to induce DNA damage in *C. elegans* (Figure 6A). This suggests that the relationship between DNA damage and microplastic concentration is not linear – once exposure exceeds the threshold required to elicit a damage response, increasing concentration does not proportionally increase the severity of DNA damage. This study also showed that smaller microplastic beads (0.2 μm) can induce a stronger DNA damage response in *C. elegans* than bigger microplastic beads (1 μm) (Figure 6B), likely because smaller microplastic particles are easier to get into *C. elegans* and their cells to induce DNA damage.

Future directions:

This study measured DNA damage indirectly using the transgenic *cep-1::GFP* *C. elegans*, which expresses GFP in response to DNA damage.²⁷ To directly quantify DNA damage at the single-cell level, future studies could employ the comet assay, which visualizes and quantifies DNA fragmentation and damage.²⁹ In addition, the long-term effects of microplastic exposure on DNA integrity and health outcomes in *C. elegans* beyond the 24-hour window used in this study warrant further investigation. Future studies could also extend this work to microplastics derived from other widely used plastic polymers, such as polyethylene and polypropylene. Finally, to assess the relevance of these findings to human health, human cell lines could be used to investigate whether microplastics induce DNA damage in a mammalian system.

Underlying mechanisms:

How do microplastics cause DNA damage? Microplastics may enter cells through endocytosis or engulfment. The accumulation of microplastics in cells can lead to DNA damage through several potential mechanisms. First, microplastics can trigger oxidative stress within cells by stimulating the excessive generation of reactive oxygen species (ROS) in response to foreign materials (microplastics). ROS can damage a wide range of cellular components, including DNA, leading to mutations or strand breaks.^{30,31} Second, microplastics can initiate inflammatory responses. Inflammatory cells produce ROS and other reactive molecules that can cause DNA damage in surrounding cells. Third, microplastics can disrupt cellular membranes, leading to cytotoxicity. This disruption can lead to the release of cellular contents or the entry of harmful substances, which can cause direct DNA damage or interfere with DNA repair processes. Finally, microplastics might interfere with the cellular machinery responsible for DNA replication, potentially causing replication errors, chromosomal aberrations, or double-stranded breaks. Collectively, the mechanisms by which microplastics induce DNA damage are likely multifactorial and may vary depending on particle size, chemical composition, concentration, exposure duration, and cellular context.

Further research is needed to fully elucidate these underlying mechanisms and their relative contributions to genotoxicity.

■ Conclusion

This study investigates the effects of polystyrene microplastics on causing DNA damage response in the multicellular model organism *C. elegans*. We show that polystyrene microplastics induce a strong DNA damage response in *cep-1::GFP* transgenic *C. elegans*, suggesting potentially similar mechanisms may operate in other multicellular organisms, including humans. Given that DNA damage underlies the genetic mutations that drive cancer and other diseases, these preliminary findings in a model organism suggest a potential link between microplastic exposure and disease pathogenesis and open new avenues for future investigation.

■ Acknowledgments

I would like to thank Dr. Peng Ji for his generous support and guidance throughout this project. I also thank the *C. elegans* research community for sharing transgenic strains via the *Caenorhabditis* Genetics Center, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440).

■ References

- Watson, J. D., & Crick, F. H. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* 171, 737-738 (1953). <https://doi.org/10.1038/171737a0>
- Stratton, M. R., Campbell, P. J., & Futreal, P. A. The cancer genome. *Nature* 458, 719-724 (2009). <https://doi.org/10.1038/nature07943>
- Guan, J., Wu, X., Zhang, J., Wang, Q., *et al.* Global research landscape on the contribution of de novo mutations to human genetic diseases over the past 20 years: bibliometric analysis. *J Neurogenet* 38, 9-18 (2024). <https://doi.org/10.1080/01677063.2024.2335171>
- Shendure, J., & Akey, J. M. The origins, determinants, and consequences of human mutations. *Science* 349, 1478-1483 (2015). <https://doi.org/10.1126/science.aaa9119>
- Martinez-Jimenez, F., Muinos, F., Gonzalez-Perez, A., Lopez-Bigas, N., *et al.* A compendium of mutational cancer driver genes. *Nat Rev Cancer* 20, 555-572 (2020). <https://doi.org/10.1038/s41568-020-0290-x>
- Kelpsiene, E., Ekvall, M., Lundqvist, M., Cedervall, T., *et al.* Review of ecotoxicological studies of widely used polystyrene nanoparticles. *Environ Sci Process Impacts* 24, 8-16 (2022). <https://doi.org/10.1039/d1em00375e>
- Lim, X. Microplastics are everywhere - but are they harmful? *Nature* 593, 22-25 (2021). <https://doi.org/10.1038/d41586-021-01143-3>
- Karbalaei, S., Hanachi, P., Walker, T. R., & Cole, M. Occurrence, sources, human health impacts and mitigation of microplastic pollution. *Environ Sci Pollut Res Int* 25, 36046-36063 (2018). <https://doi.org/10.1007/s11356-018-3508-7>
- Geyer, R., Jambeck, J. R., & Law, K. L. Production, use, and fate of all plastics ever made. *Sci Adv* 3, e1700782 (2017). <https://doi.org/10.1126/sciadv.1700782>
- Barnes, D. K., Galgani, F., Thompson, R. C., & Barlaz, M. Accumulation and fragmentation of plastic debris in global environments. *Philos Trans R Soc Lond B Biol Sci* 364, 1985-1998 (2009). <https://doi.org/10.1098/rstb.2008.0205>
- Zhang, Q., Xu, E., Li, J., Chen, Q., *et al.* A Review of Microplastics in Table Salt, Drinking Water, and Air: Direct Human Exposure. *Environ Sci Technol* 54, 3740-3751 (2020). <https://doi.org/10.1021/acs.est.9b04535>
- Ragusa, A., Notarstefano, V., Svelato, A., Belloni, A., *et al.* Raman Microspectroscopy Detection and Characterisation of Microplastics in Human Breastmilk. *Polymers (Basel)* 14 (2022). <https://doi.org/10.3390/polym14132700>
- Ragusa, A., Svelato, A., Santacroce, C., Catalano, P., *et al.* Plastics: First evidence of microplastics in human placenta. *Environ Int* 146, 106274 (2021). <https://doi.org/10.1016/j.envint.2020.106274>
- Leslie, H., A. Velzen, M. J. M., Brandsma, S., Lamoree, M. H., *et al.* Discovery and quantification of plastic particle pollution in human blood. *Environ Int* 163, 107199 (2022). <https://doi.org/10.1016/j.envint.2022.107199>
- Amato-Lourenco, L. F., Carvalho-Oliveira, R., Junior, G., Mauad, T., *et al.* Presence of airborne microplastics in human lung tissue. *J Hazard Mater* 416, 126124 (2021). <https://doi.org/10.1016/j.jhazmat.2021.126124>
- Nihart, A. J., Garcia, M., West, A., Campen, M., *et al.* Bioaccumulation of microplastics in decedent human brains. *Nat Med* (2025). <https://doi.org/10.1038/s41591-024-03453-1>
- Li, P., & Liu, J. Micro(nano)plastics in the Human Body: Sources, Occurrences, Fates, and Health Risks. *Environ Sci Technol* (2024). <https://doi.org/10.1021/acs.est.3c08902>
- Wright, S. L., & Kelly, F. J. Plastic and Human Health: A Micro Issue? *Environ Sci Technol* 51, 6634-6647 (2017). <https://doi.org/10.1021/acs.est.7b00423>
- Mattsson, K., Johnson, E., Malmendal, A., Cedervall, T., *et al.* Brain damage and behavioural disorders in fish induced by plastic nanoparticles delivered through the food chain. *Sci Rep* 7, 11452 (2017). <https://doi.org/10.1038/s41598-017-10813-0>
- Brenner, S. The genetics of *Caenorhabditis elegans*. *Genetics* 77, 71-94 (1974). <https://doi.org/10.1093/genetics/77.1.71>
- Stiernagle, T. Maintenance of *C. elegans*. *WormBook*, 1-11 (2006). <https://doi.org/10.1895/wormbook.1.101.1>
- Rodriguez, M., Snoek, L. B., De Bono, M., & Kammenga, J. E. Worms under stress: *C. elegans* stress response and its relevance to complex human disease and aging. *Trends Genet* 29, 367-374 (2013). <https://doi.org/10.1016/j.tig.2013.01.010>
- Yuan, J., Shaham, S., Ledoux, S., Ellis, H. M., & Horvitz, H. R. The *C. elegans* cell death gene *ced-3* encodes a protein similar to mammalian interleukin-1 beta-converting enzyme. *Cell* 75, 641-652 (1993). [https://doi.org/10.1016/0092-8674\(93\)90485-9](https://doi.org/10.1016/0092-8674(93)90485-9)
- Williams, P. L., Anderson, G. L., Johnstone, J. L., Wedeking, P., *et al.* *Caenorhabditis elegans* as an alternative animal species. *J Toxicol Environ Health A* 61, 641-647 (2000). <https://doi.org/10.1080/00984100050195125>
- Greiss, S., Schumacher, B., Grandien, K., Rothblatt, J., & Gartner, A. Transcriptional profiling in *C. elegans* suggests DNA damage dependent apoptosis as an ancient function of the p53 family. *BMC Genomics* 9, 334 (2008). <https://doi.org/10.1186/1471-2164-9-334>
- Liu, Z., Hua, X., Zhao, Y., Bian, Q., & Wang, D. Polyethylene nanoparticles cause reproductive toxicity associated with activation of both estrogenic hormone receptor NHR-14 and DNA damage checkpoints in *C. elegans*. *Sci Total Environ* 906, 167471 (2024). <https://doi.org/10.1016/j.scitotenv.2023.167471>
- de Souza Machado, A. A., Kloas, W., Zarfl, C., Hempel, S., & Rillig, M. C. Microplastics as an emerging threat to terrestrial ecosystems. *Glob Chang Biol* 24, 1405-1416 (2018). <https://doi.org/10.1111/gcb.14020>
- Xu, Y., Ou, Q., Jiao, M., Liu, G., & van der Hoek, J. P. Identification and Quantification of Nanoplastics in Surface Water and Groundwater by Pyrolysis Gas Chromatography-Mass Spec-

- trometry. *Environ Sci Technol* 56, 4988-4997 (2022). <https://doi.org/10.1021/acs.est.1c07377>
29. Olive, P. L., & Banath, J. P. The comet assay: a method to measure DNA damage in individual cells. *Nat Protoc* 1, 23-29 (2006). <https://doi.org/10.1038/nprot.2006.5>
30. Martins, S. G., Zilhao, R., Thorsteinsdottir, S., & Carlos, A. R. Linking Oxidative Stress and DNA Damage to Changes in the Expression of Extracellular Matrix Components. *Frontiers in genetics* 12, 673002 (2021). <https://doi.org/10.3389/fgene.2021.673002>
31. McCord, J. M. The evolution of free radicals and oxidative stress. *Am J Med* 108, 652-659 (2000). [https://doi.org/10.1016/s0002-9343\(00\)00412-5](https://doi.org/10.1016/s0002-9343(00)00412-5)

■ Author

Renee Ma is a junior at New Trier High School with a passion for biomedical research and public health.

How To Solve a Rubik's Cube: Using Search Algorithms

Richard Y. Tao

Macleans College, 2 Macleans Road, Bucklands Beach, Auckland, 2012, New Zealand; richard82288@gmail.com

ABSTRACT: Breadth-first search (BFS) and depth-first search (DFS) are 2 of the most common methods of graph traversal to get from a starting node to a goal node. This paper explores how BFS, DFS, and a combination of the two—Iterative-deepening depth-first search (IDDFS)—compare when solving a Rubik's Cube. To investigate their differences in performance and correctness empirically, a linear representation of the cube, the moves (quarter-turns as functions), and the orders of instructions for each search algorithm were developed in code. I hypothesized that IDDFS would achieve the lowest runtime of the three algorithms while preserving the optimal solution length (H1). Furthermore, BFS and IDDFS would yield optimal solutions more reliably than DFS, measured by the success rate of solutions found by the algorithm (H2). IDDFS beats BFS and DFS in runtime, yielding optimal solutions in the fastest time. The results validate both H1 and H2.

KEYWORDS: Computer Science, Graph Traversal, Search Algorithms, BFS/DFS.

■ Introduction

The Rubik's Cube is the best-selling toy and puzzle of all time, so there has been an aura surrounding its challenge.¹ Traditional solving methods, such as CFOP (Cross, first 2 layers, orienting last layer, permuting last layer), Roux, ZZ, and Petrus, are widely used in the speedcubing world today, but these generate average solutions of around 50 moves.² The Fewest Moves Challenge (FMC) was introduced to the World Cube Association (WCA), the official governing organization for all competitive speedcubing competitions, in 2015, with the challenge being as stated in its name. Competitors have 1 hour, 3 cubes, a pen, and paper to find the shortest solution they can to a given scramble. God's number is the largest number of moves it would take to solve any cube state, assuming the optimal solution. It is known to be 20 moves.³ In a hobby and sport dominated by human interest, how would computers fare with modern-day search algorithms? Is there an optimal method for finding the shortest solution to any scramble? If there is, humans haven't managed to do it yet, but maybe computer search algorithms can. Korf's, Thistlewaite's, and Kociemba's algorithms have produced promising, near-optimal results.⁴

2-dimensional search trees, featuring a starting node, a goal node, and everything in between connected by branches, are also prevalent in the modern world. Google Maps gets you from point A to point B using heuristic functions dependent on factors such as traffic, distance, and time, which all factor into the heuristic cost and can be represented by a set of nodes and branches joining these nodes.⁵ In the context of Rubik's Cube solving, although classic properties of BFS, DFS, and IDDFS are well known, their empirical trade-offs on short scrambles remain under-quantified. Accordingly, we test two hypotheses:

H1: For short scrambles (≤ 7 moves), IDDFS exhibits lower median runtime than BFS and DFS while returning optimal-length solutions

H2: BFS and IDDFS return optimal solutions at a higher rate than DFS across depths 1–7

The roots of this topic come from finding God's number for the Rubik's Cube, where an upper and lower bound were squeezed through a 3-decade process until Tomas Rokicki *et al.* used Google's supercomputers to find the number to be 20.³ Herbert Kociemba contributed to the process through Kociemba's algorithm, which essentially identifies a subset of states that can all be solved in 18 moves.⁶ Every scramble can undergo a series of moves to become a state of this subset, which can subsequently be solved.⁷ Although God's number is 20 and not 18, the most common optimal solution length to any random scramble is 18 moves,⁸ proving Kociemba's algorithm's usefulness. Kociemba's algorithm is near but not optimal,⁶ but by a combination of a known subset and considering heuristics, it can always find a near-optimal solution in a reasonable amount of time.⁹

Korf's algorithm can find the shortest solution to any Rubik's Cube scramble, but it is very computationally intensive and requires significant amounts of resources and time to execute. Subsequent research attempts have either employed new techniques to lessen the run time by about 10%,⁶ or depended on heuristics (IDA*) and known subsets to balance time and cost against finding a near-optimal solution.

My research employs 3 fundamental search algorithms – BFS, DFS, and IDDFS – to solve the cube and analyze their inherent differences in relation to the topic. These 3 methods have been chosen logically: they are the fundamental basis of most search algorithms. Despite many papers detailing the methods used to solve a cube using different search algorithms, there does not seem to be an existing direct comparison between the accuracy and runtime of these methods. Although not done at the higher move counts, which are closer to a true scramble length, due to lack of memory, the pros and flaws of each search algorithm can be easily seen for scrambles less than 8 moves, and also compared against the theoretical search space. Something which I have done that may not have been reflected on as much in other previous literature is tracking the

scrambles used and solutions produced, as paired with certain move orders; this may produce interesting results for the different search algorithms. The comparison of these 3 different search algorithms in combination with certain scrambles can help us to better understand where and when to apply these fundamental methods when solving the Rubik's Cube computationally.

■ Methods

The first step is representing the cube. A cube net was used, with a complete list of 54 numbers representing each facelet on the puzzle. There are 6 faces on a Rubik's Cube, hence 12 unique 90° turns (e.g., U for 'up' face clockwise, and U' for 'up' face counterclockwise). Each of the 12 possible quarter turns was then set as a function, each shifting 20 facelets in a certain way (1 turn moves 21 facelets, less the centerpiece, which is unaffected). In a graph with nodes representing the states of the cube and edges connecting these states via quarter-turn moves, BFS is a sweeping algorithm, exploring every possible move at each depth, then moving on to the next. This ensures the shortest solution is found each time due to the algorithm's completeness. With big O notation, where $O(\cdot)$ represents the most amount of time required in the worst case scenario to reach the goal node, the time complexity (big O) of BFS is $O(bd)$, where b represents the branching factor of our graph (how many children from each node, in this case, 12), and d being the depth of the goal node; the space complexity is $O(bd)$ as every node explored at each level is kept in memory.¹⁰ BFS employs a FIFO (first in, first out) queue, meaning each depth is processed from left to right in a designated move order, guaranteeing the shortest solution each time. It should also be noted that the general order of action for each algorithm is: check if goal state → explore next node → repeat, in case the initial state is indeed already the goal state.

DFS employs LIFO (last in, first out) with a stack, adding to the stack by exploring the deepest paths first, then returning to the last depth level after hitting a dead end. It has a time complexity of $O(bd)$ and space complexity of $O(bd)$.¹¹ This space complexity is much less than BFS, increasing linearly rather than exponentially with depth, as DFS only remembers the path it took to get to a node, along with all the daughters of the nodes along that path. Immediate inverse moves are pruned trivially, thus saving time visiting unnecessary paths and preventing endless loops. For experimental reasons, the depth limit was set to 7, meaning that our DFS algorithm would only return solutions for scrambles less than or equal to 7 moves in length. A higher depth limit would also likely increase solve times for lower move counts, so it is only fair to set it to be the same throughout the entire experiment.

IDDFS is a combination of BFS up to a certain depth, but expanding its depth limit by 1 after each level, as the name suggests. It is effectively DFS being iterated at increasing depths (as the name suggests). The time complexity of IDDFS is the same as for BFS and DFS individually ($O(bd)$), as in the worst-case scenario, every node will have to be visited before finding the goal node. Immediate inverse moves are pruned just like in DFS, which plays out much better for time: due

to IDDFS's complete, breadth-first character, the branching factor is reduced from 12 to 11, greatly reducing its realistic search space. Its space complexity is the same as DFS ($O(bd)$) due to the algorithm's intrinsic depth-first nature. Effectively, IDDFS is an extension of DFS, which guarantees the completeness of BFS and the ability to find the optimal solution each time while maintaining the superior space complexity of DFS, making it theoretically the ideal search algorithm for solving the Rubik's Cube in our case.

A fixed moveset is also specified for every search algorithm to define the search space and thus the order of execution with FIFO and LIFO queues. The move set is U, U', D, D', L, L', R, R', F, F', B, B'. It is arbitrary, but it may create varying search times for the greedy DFS depending on what moves are in the solution.

Here is the link to the code, which is contained in the GitHub repository:

https://github.com/richard82288-del/rubiks_cube_search_algo

■ Results and Discussion

Results:

https://docs.google.com/spreadsheets/d/1iqReiS-QxsAv-I5cNupY_3d03f4BrLx1C4o7KK-S7tzc/edit?usp=sharing

All 3 algorithms employed in this experiment are uniform-cost (non-cost-associated), as each move on a Rubik's Cube objectively gives an equal probability of reaching the goal state when the solution is unknown. While BFS gives the optimal solution each time as expected, DFS sometimes adds unnecessary moves to the solution, highlighting its greed for depth. This often results in an equivalent 'U3' instead of 'U', or adding an 'F', 'F''; as its search order prioritizes clockwise moves before anticlockwise moves, DFS occasionally will add 2 moves to otherwise optimal solutions even in the <6 move range. However, given lucky move sets often involving sequences of U's, DFS may occasionally be faster than expected, leading to a rather unpredictable time-to-move ratio, although rarely outperforming BFS or IDDFS in time or solution length (solution length is rather redundant when analyzing shorter scrambles, as there is usually only 1 solution from an algorithmic perspective).

BFS crashes the system at 7 move scrambles, as my 8GB of RAM is not enough to sustain its large but complete space complexity. Its requirement for great operational capacity or memory, especially for a problem such as the Rubik's Cube, which has a branching factor of 12 (in our case, ignoring double moves), makes it unfeasible for slightly longer and more realistic scrambles.

The DFS stack significantly mitigates BFS's memory problem, narrowing down its path to a solution. This does, however, often lead DFS down the wrong path, often taking longer to find depth 1 solutions than for higher depth scrambles. Its inability to accept a dead end on a puzzle with no such thing as the Rubik's Cube highlights its weakness with greed for depth, making it very unreliable for solving a Rubik's Cube, although

much more likely to find a solution at greater depths than BFS due to its tiny memory footprint.

IDDFS typically provides the optimal solution with the fastest time out of all three algorithms, optimal due to its complete nature and faster than BFS because of its small memory.

Table 1: Number of nodes (states) in the search space for a given scramble length. The search space increases exponentially, as can be seen in the table. Note the branching factor is 11 to remove immediate inverse moves, and the numbers are still rough as there are looping paths, e.g., U U U instead of U', which appears in DFS solutions rather frequently.

# of moves in scramble	N	1	2	3	4	5	6	7	8
Search space (# of states)	$\frac{n(1-r^n)}{1-r}$	12	144	1,596	17,568	193,260	2,125,872	23,384,604	257,230,656

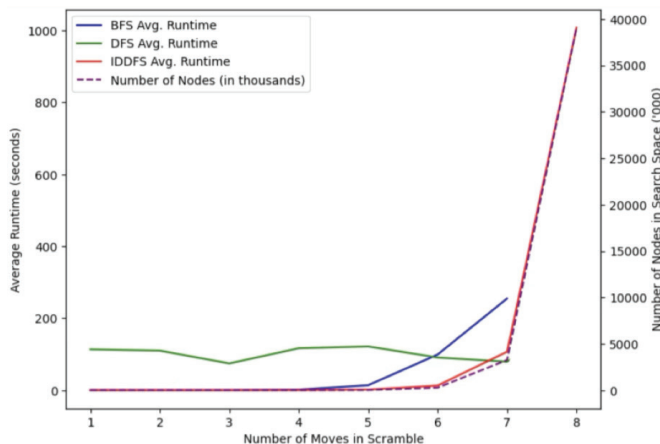


Figure 1: Average runtimes across the number of moves in the scramble for BFS, DFS, and IDDFS. The time taken for IDDFS to find a solution for n moves closely follows the trend for the number of nodes in the search space. BFS runs out of space at 8 moves, and DFS seems to take a similar amount of time regardless of move count.

As can be seen in Figure 1 above, comparing the number of moves in the scramble with the number of nodes in the search space (Table 1) and average runtime for each search algorithm, the number of nodes in the search space increases exponentially. IDDFS also has a trend similar to this increase in the number of nodes, as its systematic iterative deepening at each level is dependent on the length of the scramble (and solution), therefore searching every node at each depth with linear memory, so it does not store every path. DFS takes about the same time to return a solution regardless of scramble length, which makes sense due to its nature of searching deep solutions and potentially missing shorter ones. Note that the depth limit for DFS is set to be 7 moves due to computational constraints. The average time taken by DFS to return no solution for 8 moves is also, in essence, the time taken for it to visit all nodes up to 7 moves, which is actually faster than the average time taken by BFS to generate a solution for 7 moves, but still slower than IDDFS (Figure 1). In further experiments, god's number, the fact that any state of the cube can be solved in 20 moves,³ may be used to set the depth limit, but this may change the runtimes for lower move counts for DFS, as its search space is greatly expanded. The more efficient method for setting DFS's depth limit would be to use the scramble length, but this is unrealistic for solving unknown scrambles.

Although IDDFS is typically the fastest, DFS's reliance on luck is evident with one scramble "U, U", where it found the solution "U, U" in by far the fastest time for any 2-mover (Results). This is expected as U is the first move in the set move order, and it just so happens that in this case, the first path explored by DFS is the solution, while both BFS and IDDFS have to sweep the first depth. The runtime for DFS seems to be random, as, due to its inherent nature, DFS is largely dependent on the specific moves that make up the solution.

Conclusion

BFS works well for solving a Rubik's Cube when there is plenty of RAM available, and especially for shorter scrambles where it is able to guarantee the optimal solution in a reasonable amount of time. DFS has the advantage of a small memory footprint, but its reliance on a known and set depth limit makes it both unreliable and unrealistic for solving a Rubik's Cube. Even with God's number of 20, DFS may produce unnecessarily long solutions or fail to find a solution even after a long time, as it is luck-based. However, when the set depth level is similar to the scramble/solution length, DFS may even be faster than IDDFS.

A combination of BFS and DFS, in essence, we can conclude that IDDFS is the best search algorithm of the three for solving a Rubik's Cube. With the optimality of BFS and the scalability of DFS, IDDFS's iterative deepening can guarantee the optimal solution in a reasonable amount of time using a reasonable space complexity.

Given enough RAM, further experimentation can be done to discover the limits of IDDFS with respect to Rubik's Cube solving and how it compares with IDA* in terms of time taken to find the optimal solution. Setting DFS's depth limit to god's number, 20, would also be worthwhile to see its results. I have not done so for the sake of time, and the main goal is to compare these 3 search algorithms rather than delving deep into 1.

The Rubik's Cube is just one example of the applications of a search algorithm, which I happen to be very interested in. A better understanding of the risks and benefits involved with each of these fundamental search algorithms in graph theory helps us to make better decisions when it comes to similar real-life situations where there are legitimate consequences.

Acknowledgments

I would like to acknowledge Qi Xuan Khoo for being my mentor for this research article.

References

1. Daniels, J. S., Moreau, D., & Macnamara, B. N. Learning and Transfer in Problem-Solving Progressions. *Journal of Intelligence*, 10(4), 85. <https://doi.org/10.3390/jintelligence10040085>. 2022.
2. Duberg, D., & Tidestrom, J. Comparison of Rubik's Cube Solving Methods Made for Humans. <https://www.diva-portal.org/smash/get/diva2:812006/FULLTEXT01.pdf>. 2015.
3. Rokicki, T., Kociemba, H., Davidson, M., & Dethridge, J. The diameter of the Rubik's Cube group is twenty. *siam REVIEW*, 56(4), 645-670. <https://tomas.rokicki.com/rubik20.pdf>. 2014.

4. Chen, J. Different Algorithms to Solve a Rubik's Cube (J. Phys.: Conf. Ser. 2386 012018). <https://iopscience.iop.org/article/10.1088/1742-6596/2386/1/012018/pdf>. 2022.
5. Mehta, H., Kanani, P., & Lande, P. Google Maps (International Journal of Computer Applications (0975 – 8887) Volume 178 – No. 8). https://www.researchgate.net/profile/Pratik-Kanani/publication/333117435_Google_Maps/links/5eb6cc7da6fdcc1f1dcb10aa/Google-Maps.pdf. 2019, May.
6. Grumser, T. Engineering Optimal Solvers for Rubik's Cubes (Doctoral dissertation, Karlsruher Institut für Technologie (KIT)). <https://ae.iti.kit.edu/download/grumser-tan.pdf>. 2022.
7. Rokicki, T. Twenty-Five Moves Suffice for Rubik's Cube (No. arXiv:0803.3435). arXiv. <https://doi.org/10.48550/arXiv.0803.3435>. 2008.
8. Korf, R. E. Finding optimal solutions to Rubik's Cube using pattern databases. In AAAI/IAAI (pp. 700-705). <https://cdn.aaai.org/AAAI/1997/AAAI97-109.pdf>. 1997, July.
9. Kaur, H. Algorithms for solving the Rubik's Cube (KTH Royal Institute of Technology). <https://www.diva-portal.org/smash/get/diva2:816583/FULLTEXT01.pdf>. 2015, May 27.
10. Iterative Deepening Search(IDS) or Iterative Deepening Depth First Search(IDDFS). GeeksforGeeks. <https://www.geeksforgeeks.org/dsa/iterative-deepening-searchids-iterative-deepening-depth-first-searchiddfs/>
11. Lozano-Pérez, T., & Kaelbling, L. P. Notes on the Complexity of Search (MIT 6.034 Artificial Intelligence, Fall 2003). Massachusetts Institute of Technology. <http://www.ai.mit.edu/courses/6.034b/searchcomplex.pdf>. 2003, September 23.

■ Author

Richard Tao is a Rubik's puzzle enthusiast who is a senior at Macleans College in Auckland, New Zealand. He is the founder of Cubing East Auckland, a non-profit organization promoting speedcubing in the local community. Richard plans to go to college in the United States one day, majoring in mathematics.

Efficiency of Lipid Nanoparticles in Glioblastoma Therapies

Saimanasaa Viswanathan

Emerson High School, 6300 Collin McKinney Pkwy, McKinney, TX 75070, USA; saimanasaa.viswanathan@gmail.com

ABSTRACT: Increasingly targeting the older population, Glioblastoma is a grade four brain tumor with less than a 10% survival rate. One major challenge in treatment is delivering drugs across the Blood-Brain Barrier (BBB) and Blood Tumor Barrier (BTB). Chemotherapy has been used as a main solution for fighting tumors, specifically in Glioblastoma; however, it can cause many side effects for patients, such as headaches and changes to mental function. Nanoparticles, and particularly lipid nanoparticles, have emerged as leading drug delivery methods to overcome those challenges. Due to their small size, they can effectively cross the BBB and can be modified with other targeted ligands. Three types of lipid nanoparticles are discussed: Solid Lipid Nanoparticles (SLNs), Liposomes, and Nanostructured Lipid Carriers (NLCs). With their liquid lipids and flexible crystallized structure, NLCs can help overcome major shortcomings of chemotherapy and improve future outcomes for Glioblastoma patients.

KEYWORDS: Cellular and Molecular Biology, Neurobiology, Glioblastoma, Blood-Brain Barrier, Lipid Nanoparticles.

■ Introduction

The impacts of cancer are everlasting and can be devastating for the body, particularly in the case of brain tumors. They can develop as a result of external factors, such as UV rays from the sun, or through genetic errors that occur during mitosis, when cells divide to create new cells. The stages of this process include:

- Interphase: the cell grows and replicates its DNA
- Prophase: chromosomes become visible
- Metaphase: chromosomes align to the center of the cell
- Anaphase: sister chromatids separate
- Telophase: cells prepare to divide

The cells cannot move on to the next stage without stopping at checkpoints. These make sure the cells go through the respective stage properly. When the mutated DNA proceeds into the cell cycle, it can bypass the checkpoints and replicate uncontrollably. The mutations created can lead to gain-of-function (GOF) phenotypes, which promote the altered genes to enhance tumor growth and formation. This action can result in tumors.

In Glioblastoma, a type of brain tumor, patients only have a 5-10% survival rate with standard treatment.¹ Because the brain is responsible for many integral functions, such as storing memory, cognitive thinking, processing emotions, movement, and regulating the body, it is crucial to implement treatments that not only effectively dissipate the tumors but also mitigate the threatening effects associated with them. Chemotherapy treatments, for instance, are often used. However, not only do they target the tumor cells, but also the healthy cells, as part of their non-selective process. Chemotherapy is inefficient because it cannot differentiate between healthy and unhealthy cells.

Due to this problem, loss of important red and white blood cells can be significant when transporting medicine through chemotherapy, which can produce side effects such as fatigue and hair loss.² Nanomedicine, however, can selectively target

only tumor cells, reducing these side effects tremendously. Within the branch of nanomedicine, nanoparticles are an important tool used especially for targeted drug delivery of medication. With the efficient applications of nanomedicine, resolving brain tumors will be associated with less stress for patients during treatments.

Nanomedicine-based treatments indeed result in improved reduction of brain tumors. According to Kumar *et al.*,³ polymeric nanoparticles have provided many positive responses in treating brain tumors. This is due to their properties of having high biodegradability, permeation in tumor cells, and a surface-to-volume ratio. The high permeation efficiency allows cells to produce further connected pore spaces, which can then allow for liquids to flow efficiently. This is useful, especially when implementing a drug-delivery system, which is a technology used to carry drugs throughout the body. These are a must for cancer therapy treatments, as that is the way to carry medicine into the brain. However, the problem arises when the drugs have to pass through the BBB (Blood-Brain Barrier) and BTB (Blood-Tumor Barrier) Complex, both of which are highly semipermeable barriers in the brain. The BBB only allows important materials into the brain, such as nutrients and oxygen, while turning away others, like pharmaceuticals. The BTB is a barrier formed due to the internal pressure within the tumor setting. However, nanoparticles (NPs) can be effectively used to cross both of the barriers. Along with their selective capabilities, they can transport multiple drugs throughout the brain, utilize many signaling pathways, and can effectively be transported across these barriers.⁴ There can be a positive correlation seen between the use of NPs and the healing of brain tumors, as the particles are effectively able to go into the brain and deliver medications successfully.

Further, nanoparticles can take advantage of the glioma EPR effect in the cell.⁵ The EPR effect is when nanoparticles can accumulate within tumor cells, and this is efficient as they will be able to target and dissipate them efficiently. Meola and

Rao emphasize the importance of their small size and shape as a beneficial contributor to delivering treatments.⁶ This property of nanoparticles helps to improve efficiency in tumor treatments, especially in the brain, where there are many small pores and barriers for medications to cross. However, Glioblastoma often retains some parts of the BBB, limiting nanoparticles from getting through. As a result, the EPR effect may not always be sufficient for lipid nanoparticles to perform drug delivery.

Additionally, a study conducted utilized the unique functionalized surface of the NPs and attached antibodies and a chemotherapy drug to them. According to the data, the NPs helped to increase the survival rate of the mouse with Glioblastoma by 50%.⁷ Researchers stated it is a crucial finding since the largest problem in medicine initially was trying to pass the chemotherapy treatments into the brain. Still, with this now, it's indeed possible.

Given the confirmed correlation between the use of nanomedicine and the efficient healing of brain tumors, NPs and nanotechnology must be implemented in treatments. The brain is one of the most important organs carrying out a multitude of vital functions; the body simply cannot function without it. Not only will NPs increase the effectiveness of delivering medication, but they will also not interfere with a patient's brain health or activities.

■ Methodology

The primary goal of this research was to determine which type of lipid nanoparticle is most efficient and compatible for treating the Glioblastoma brain tumor. This research was completed through a secondary literature review, which aims to compile multiple academic journals and primary sources, both for health and effectiveness. Studies published between 2016 and 2025 were included to capture the most recent advancements of lipid nanoparticle treatments for Glioblastoma. This time frame was selected because many recent advancements have been made, making earlier studies less representative of current data. The literature search was conducted using keywords such as “blood-brain Barrier,” “lipid nanoparticles for treatment,” and “Glioblastoma brain tumors.” Data were collected to analyze nanoparticles’ efficiency in reaching the brain and successfully treating the tumors through drug-delivery systems. The sources referenced include scientific papers published in peer-reviewed journals, clinical data, and other academic literature globally. Therefore, this literature review is derived from multiple backgrounds, leaving a well-organized, coherent review for accessing research on this specific area of study. Further research on diagnostics and other applications of lipid nanoparticles was also included. This study did not conduct experiments. Rather, it compiled and discussed data presented in the primary sources and academic literature.

The Glioblastoma Brain Tumor:

Glioblastoma is known to be one of the most common and dangerous types of brain tumors. Even with its devastating impact on the brain, there is no specific medicine known to cure the tumor completely.⁸ This aspect is crucial as the brain con-

ducts a multitude of functions, not limited to: enacting speech, movement, memory, and regulating breathing and temperature. Glioblastoma can eventually prevent the brain from performing its daily functions, not only impacting its ability to live at a higher standard but even at the cost of death.

The origins of tumor formation start in astrocytes, which are a type of glial cell. These cells support nerve cells, the spinal cord, and the brain.⁹ Astrocytes help to protect the neurons against any unneeded or toxic compounds, as well as form synapse connections within the brain. However, with glioblastoma forming from them, it can break their functions and no longer be effective. It can then go on to damage and eventually destroy the healthy cells and tissues in either the brain or the spinal cord, depending on where it started.⁸ It is possible for anyone of any age to be diagnosed with Glioblastoma, although it's more common to be present in adults, with the median age being 64 and the peak age being 75-84 years old.¹⁰ Due to the body's aging, cellular processes like genetic mutations are more likely to lose control, as well.

Glioma is a type of tumor that occurs in the brain and is categorized into two main sections: low-grade and high-grade gliomas. Low grade comprises grades 1 and 2, and high grade comprises grades 3 and 4; the order is in increments of aggressiveness. Glioblastoma is a type of glioma that is categorized as grade 4, meaning it's the most aggressive on the scale. It contains features of microvascular proliferation and necrosis (MVP), which simply emphasizes the rapidly growing glioma. Grade 4 gliomas have an average of a 5-10% survival rate, unlike grade 1, for instance, which has a 95% survival rate. The most aggressive grade 4 is the Glioblastoma *IDH* wildtype.¹¹ The effects of Glioblastoma further prove how much danger it can cause not only to the frontal and temporal lobes of the brain, where it primarily occurs, but also impacts the overall well-being of the body (Figure 1).

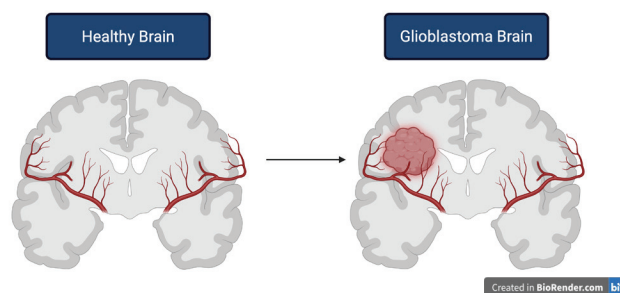


Figure 1: Healthy Brain vs. Glioblastoma Brain. This image illustrates the difference between a healthy brain and one diagnosed with Glioblastoma; the brain with Glioblastoma has a tumor present, which can affect various aspects of brain functioning. These are not limited to tumor growth, which can cause an increase in intracranial pressure due to weight distribution, as well as the breakdown of the blood-brain barrier. These chemical side effects, which will be discussed later in detail, are crucial components of the dangers of Glioblastoma on the brain, created using BioRender by Saimanasaa Viswanathan.

Causes and Symptoms:

Although there isn't a specific root cause to this tumor, mutations can be integral to contributing to the tumor's existence. Mutations happen through a single change or multiple changes occurring in a DNA sequence.¹² The DNA sequence is important as DNA communicates to the cell what to do.

However, in cancer cells, the cells will enact the mutated DNA to keep replicating; this multiplication of cancer cells, constantly replicating, forms tumors, while the healthy cells start to die off in response to the large amount of cancer cells. For instance, in the frameshift mutation, a base can either be deleted or added to the sequence. However, this can alter the entire base pairing, as proteins read bases in groups of three, known as codons. When the order shifts, different proteins will be produced. Other types of shifts include point mutations, which only affect a single nucleotide; for example, one base can be altered for another base, meaning that the entire sequence won't change. Furthermore, there are also missense and nonsense mutations, which also change the order of the DNA sequence.

Some of the key mutations present in Glioblastoma most commonly include *TP53* and/or *PTEN*. According to Olafson *et al.*, the *TP53* gene can acquire the GOF phenotypes (gain of function), which can aggravate tumor growth. Furthermore, this gene being mutated can result in a p53 altered protein, which is unable to activate apoptosis in the cells, i.e., the action of cell self-destruction.¹³ Apoptosis is important because if it doesn't happen, then the mutated or damaged DNA will continue to replicate and spread, elevating growth into a tumor.¹⁴ The *PTEN* gene contains the PTEN enzyme. PTEN mutations reduce tumor suppression activity, leading to uncontrolled cell growth. The symptoms of Glioblastoma are shown (Figure 2).

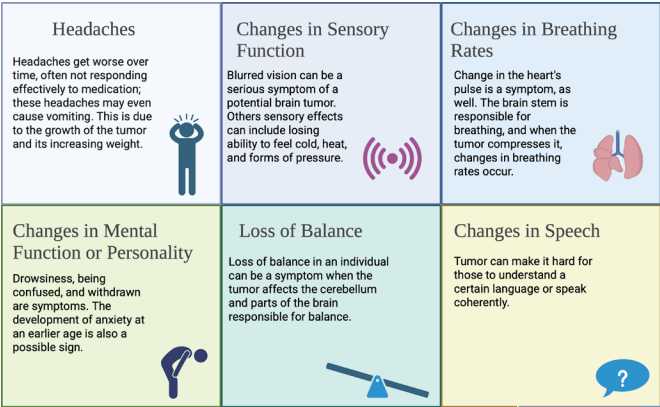


Figure 2: Symptoms of Glioblastoma in the Human Body. The symptoms of Glioblastoma start within the brain first and send signals for further action; these signs are crucial for engaging in medical treatment. Some of these are headaches, sensory changes, changes in breathing rates, changes in medical function or personality, loss of balance, and changes in speech. Information gathered from <https://www.mdanderson.org/cancer-types/glioblastoma/glioblastoma-a-symptoms.html> and created using BioRender.

Challenges to Chemotherapy:

Chemotherapy is a commonly used treatment for Glioblastoma and other cancers. However, along with its usage, multiple drawbacks prevent it from being efficient in comparison to other medications. The major consequence of chemotherapy is its inability to distinguish between cancer cells and healthy cells. Oftentimes, it is prone to targeting both cancer and healthy cells, resulting in major side effects for patients such as weight changes, fertility barriers, constipation, and appetite changes.² Its limitations are not only with these

effects but also with its ability to deliver the drugs to the body efficiently. Additionally, the size of chemotherapy drugs makes it difficult to pass medications through smaller areas such as the blood-brain barrier and blood-tumor barrier, contributing to less efficiency in transporting medication.

Blood-Brain Barrier:

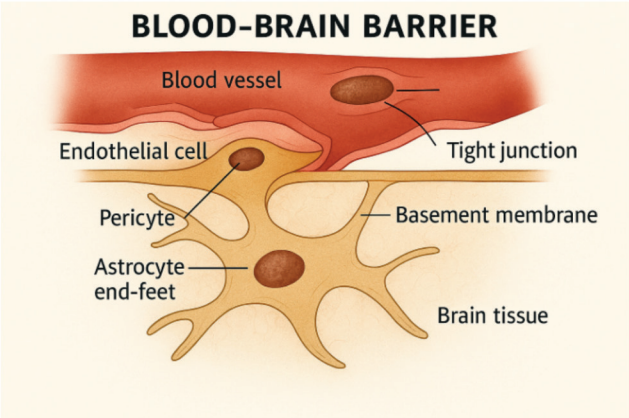


Figure 3: Blood-brain Barrier Structure. The endothelial cells form tight junctions that limit the flow of any hydrophilic substances from passing through the barrier and reaching the brain. The astrocyte endfeet help to convey signals to form the tight junctions of the blood-brain barrier. Likewise, the vascular BM (vascular basement membrane) is especially integral for being a physical support in maintaining the brain's shape as well as assisting in regulating substances. The components of the blood-brain barrier contribute to its selectivity, only allowing for particles with certain characteristics to get through, and are created using CoPilot AI.

Formation and Origin:

The blood-brain barrier (Figure 3) is a selective, semipermeable barrier containing blood vessels that regulate movement between the blood and the brain.¹⁵ It is located between the blood and extracellular space present in all regions of the brain except for the circumventricular organs (CVOs).

Implications in Treatment:

Although the blood-brain barrier can be beneficial in preventing any toxins or unneeded substances from entering the central nervous system, it can create a barrier to many drug delivery systems wanting to transport any needed biopharmaceuticals or chemical drugs for greater health purposes. According to Wu *et al.*, the blood-brain barrier excludes more than approximately 98% of small-molecule-based drugs as well as 100% of macromolecular drugs, challenging the way for these substances to pass through.¹⁶

Blood-Tumor Barrier:

Formation and Origin:

The blood-tumor barrier is a structure modified from the blood-brain barrier, specifically formed during severe tumors like Glioblastoma. The severity of the tumors can result in damage to the blood-brain barrier, resulting in the blood-tumor barrier, which is a modified form of the blood-brain barrier.¹⁷

Structural and Functional Differences Compared to the BBB:

Compared to the blood-brain barrier, the blood-tumor barrier has heterogeneous permeability, meaning the ability for substances to pass in and out varies across different locations within the tumor. This results in the inability for medications to pass through sufficiently for all locations within the tumor, as the degree of effective permeability will vary; however, compared to the blood-brain barrier, it will be more flexible in the passage of substances in and out of the membrane.

Transports and Permeability:

According to Arvanitis *et al.*, due to its heterogeneous permeability, the blood-tumor barrier also presents a challenge of increasing interstitial and intracranial fluid pressure.¹⁸ Because this barrier is constrictive compared to the blood-brain barrier, the ability for fluids to escape into the interstitial space is likely. Additionally, proteins can also pass through this space with the fluids, increasing interstitial pressure. Since this interstitial space is within the tumor itself, it will present a difficulty for such treatments to go past it with the fluids and proteins present. Likewise, intracranial fluid pressure is increased due to an increase in fluid retention and tumor mass accumulated within the skull. Because the brain skull itself cannot expand, as the tumor expands and more weight is being held upon it, the intracranial fluid pressure increases; this will present an obstacle for medical treatments to pass through to the tumors.

Implications of Treatment:

Although the blood-tumor barrier presents a series of challenges on its own, given its complexities and unevenness for medications to effectively bypass it, chemotherapy treatments still struggle to pass through it. The large size of chemotherapy treatments will be a hindrance to the small, compact, and rigid spaces within the barrier. This reasoning also advocates the need for lipid nanoparticles to be able to effectively pass through the BTB, deliver medication successfully, and minimize harm through treatment.

Drug Delivery and Resistance:

Chemotherapeutic drugs can damage healthy tissues as well, which can impact how intensive the treatment can be. Additionally, because the targeting characteristic is not present in chemotherapy drugs, cancer cells are more prone to being resistant to the drug and can advance to multi-drug resistance (MDR).¹⁹ Through this advancement, heterogeneous cancer cells adapt and grow through chemotherapeutic drugs without allowing them to be a barrier to their growth. MDR has been known for cancer cells to be able to survive a variety of anti-cancer drugs, resulting in cancer cells being mildly affected or not affected at all by the treatment due to their ability to resist such medications.²⁰

Alternatives: Why Lipid Nanoparticles?:

Nanoparticles have emerged as a widespread, efficient tool to pass medications through brain tumors to reach allocated spots effectively. Unlike chemotherapy treatments, they have targeted therapy. They can pass through the difficulties that

come with the blood-brain barrier and blood-tumor barrier, one of the most difficult pathways for drugs to get through, which chemotherapy lacks.²¹ Additionally, they are size-effective, ranging from 50-200 nm.²² Three types of lipid nanoparticles will be discussed for efficient Glioblastoma treatments: Solid Lipid Nanoparticles (SLNs), Liposomes, and Nanostructured Lipid Carriers (NLCs).

Solid Lipid Nanoparticles:

Solid Lipid Nanoparticles (SLNs) have been a promising lipid nanoparticle for the treatment of Glioblastoma brain tumors. It has managed to overcome the drawbacks that particles like liposomes, emulsions, and lipids face, such as low stability and loading capacity; through a large capacity of ligands, SLNs can deliver medication and treatments by targeting specific areas, making it widely useful when treating tumors and cancer-induced objects within the brain.²³ Having a targeting system is crucial when dealing with tumors, as the prevention of damaging healthy cells is mostly protected. The SLN's structure is shown in Figure 4.

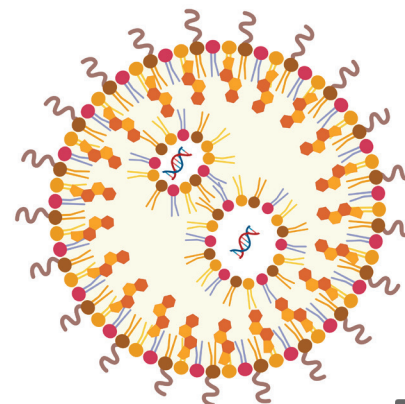


Figure 4: Solid Lipid Nanoparticles Structure. SLNs consist of solid lipids. This helps to give them a more intact, crystallized structure throughout. Drug molecules are present within. As they are lipids, they include a hydrophobic tail and a hydrophilic head, created using BioRender.

SLNs (Figure 4) have also been used to carry a variety of drugs through the drug-delivery process of medication. They can carry those with a high degree of polarity, like 5-fluorouracil, to large lipophilic molecules like taxanes; additionally, oxaliplatin and cisplatin, chemotherapeutic water-soluble drugs, are also effectively able to be loaded within SLNs.²⁵ The high versatility of these nanoparticles maximizes the potential for SLNs to transport anti-cancer drugs past the blood-brain barrier and into the targeted destination within the brain.

While SLNs have a wide spectrum of advantages and efficiency in drug delivery through Glioblastoma treatments, they tend to harden during recrystallization. This ability for these solid lipid nanoparticles to cluster together introduces an obstacle for the drugs allocated aside, making it harder for them to have space; thus, drug expulsion occurs, which is the result when there is no space left for the drug.²³ Because of their perfectly aligned, tightly packed structure, they have low drug loading efficiency, allowing space for drug molecules to fit in.²⁵

However, SLNs remain a strong candidate for Glioblastoma treatments; they are able to have high stability in water-like

environments, high drug load to carry for treatments, and can release the drug in a regulated way over the span of several weeks, instead of all at the same time.²⁶ This is important because patients receiving the treatment will not have a higher risk of side effects compared to if the drug were delivered all at once. The slow, regulated pace over this time period will not allow the body to foresee the potential adversity of side effects provided through the new medication immediately. Additionally, this improves the patients' needs more deeply. Instead of having to take daily medication or pills for treating Glioblastoma, the drugs are administered over the span of weeks, allowing the drugs to treat the patient constantly over time.

Liposomes:

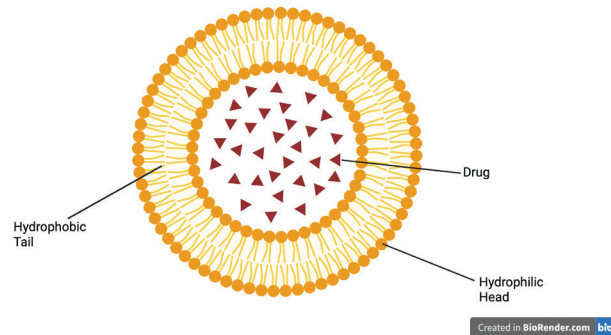


Figure 5: Liposome Structure. This figure shows how liposomes are made up of both hydrophilic heads and hydrophobic tails. They are also able to carry lipid-soluble solutions, making them efficient for the process of delivering the drugs; drugs are contained within the liposome and created using Biorender.

Liposomes (Figure 5) are another lipid nanoparticle known for their use in medical applications like drug delivery. Composed mainly of phospholipids, liposomes can either be classified as part of sphingomyelins or glycerophospholipids. There are two types of ways liposomes can load medications for drug delivery: passive or active loading methods. Within the passive loading method of water-soluble drugs, water is dissolved in the solution to form the respective liposome; likewise, in the passive loading method of hydrophobic drugs, those not compatible with water, drugs are dissolved within the lipid part of the carbon-based solvent. Then, the carbon is later evaporated, leaving just the lipid bilayers. The carbon must be evaporated because it isn't safe to keep inside the body; thus, evaporating it will leave it as the better choice. On the other hand, during active loading methods, the liposome is formed without the drug. However, for the medicine to diffuse through the lipid bilayer, mechanisms like the pH gradient technique are required for it to get inside the liposome, as a medication. Additionally, stimuli such as temperature, ultrasound, and pH have been introduced to enhance the process of drug release.²⁷ Through these processes, medications can perform in action via liposomes or lipid nanoparticles.

Liposomes contain efficient functions such as controlling the release of drugs and protecting substances from degradation. Through these passive and active methods, liposomes are selectively able to reach their highest potential dose, making it an efficient therapeutic process. However, along with their efficient physiological characteristics, they also impose some

risks regarding both their chemical compounds and manufacturing ability.²⁸

To make sure liposomes last longer in the blood, PEG chains (glycol) were introduced. Although they provide a benefit for liposomes to sustain, they introduce the PEG dilemma, which creates a barrier for liposomes to stick onto cancer cells and target them; this then makes it difficult for the liposomes to release the drug at the specific tumor sites. Although liposomes have low toxicity, high biocompatibility, and biodegradability, they come with some disadvantages: the phospholipids can suffer from lower stability and low encapsulation efficiency.²³ Additionally, the production of liposomes can be costly, and industrialization methods can be complex; this can be a challenge to process them, requiring specific equipment and methods of manufacturing.²⁹ The advantages and disadvantages of liposomes are shown (Table 1). Main methods of synthesis of liposomes for Glioblastoma tumors include Thin-Film Hydration, Reverse Phase Evaporation, and Microfluidics.

Table 1: Advantages and disadvantages of liposomes.

Advantages	Disadvantages
Encapsulation	Requires specific degrees of pH levels specifically during active loading, which increases time on waiting for the right amount.
Stability	Has shorter time frames on shelves and will need more harder storage alternatives like cold storage.
Scale-Up	Transitioning liposome production to industry is more complicated as a batch-to-batch consistency and accuracy is needed.

Nanostructured Lipid Carriers:

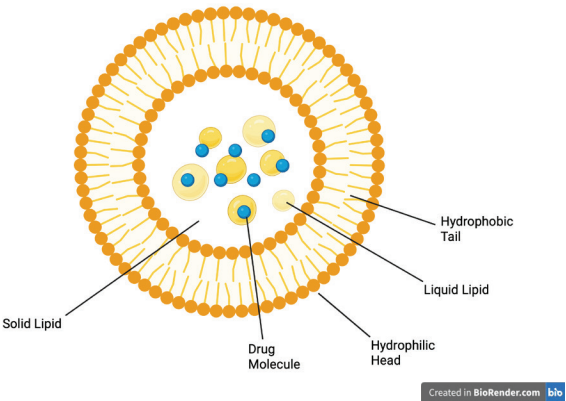


Figure 6: Nanostructured Lipid Carriers Structure. NLCs are composed of both solid and liquid lipids. The integration of liquid lipids allows the benefit of flexibility for more drug intake and carrying, giving more space for drugs to stay within the nanoparticle. As they are lipid nanoparticles, they consist of a hydrophobic tail and a hydrophilic head and were created using Biorender.

Nanostructured Lipid Carriers (NLCs) (Figure 6) are the second generation of lipid nanocarriers, having a mixture of both solid and liquid types. Because they have an imperfect crystal structure, they can have a greater efficiency in drug loading capacity; thus, because there is more space for drug delivery, there is a lower likelihood of drug expulsion.²⁵ The liquid lipids present make it possible for this flexibility. Liquid lipids found can include Capmul[®] MCM C8, sesame oil, oleic acid, and L-phosphatidyl choline (PC).²⁶ According to Xu *et al.*, NLCs have been shown in studies to not only show efficiency in drug delivery but also *in vivo* and *in vitro*. NLC

studies *in vivo*, or within an organism, show their effective nature in therapeutic efficacy and biodistribution within living organisms or the human body. NLCs are shown to be able to reach the tumor areas in need of drug delivery, as well as being effective in treating the disease. Further, they have been effective *in vitro*, in other words, studies performed outside of the organism.³⁰ As mentioned, they have a high encapsulation efficiency due to their structure, and their particle size and characteristics are suitable for stability and cellular uptake. The mixture of both solid and liquid lipids within makes it possible for high drug loading and effective measures upon transporting medication to tumors.

Additionally, studies have been made on the effectiveness of NLCs. Glioblastoma secretes many enzymes known as Matrix Metalloproteinases (MMPs). MMPs break down surrounding tissue so that the Glioblastoma tumors can spread. According to Horta *et al.*, special targeting parts were added to the NLCs, and versions were made. The version of FMBat effectively slowed down the growth of MMPs in Cancer cells and managed to cross through the blood-brain barrier.³¹ This looked promising for approaching the management of tumors. NLCs are also able to effectively carry various types of ligands such as proteins, peptides, and small molecules.³² The high ability and potential of NLCs make it possible for tumor treatments to be both efficient and promising for the needs of patients.

Although NLCs are promising treatments, they still have some limitations. Just like Liposomes, they still have to be moved to large-scale manufacturing. Keeping it consistent can be tough. Further, while they are efficient, crossing the BBB can still be a slight challenge in some circumstances due to the tight junctions present in the BBB.

Most Effective Lipid Nanoparticle:

Table 2: Comparison of Lipid Nanoparticles.

Nanoparticle	Drug Loading Capacity	BBB Penetration Efficiency	Stability	Toxicity	Scalability	Cost
Lipid Nanoparticle	Can carry hydrophilic and lipophilic drugs but moderate carrying capacity	Moderate efficiency	Relatively low- more likely to have drug leakage	Due to phospholipid composition, has low toxicity	Moderate due to complexities of lipid nanoparticles	Moderate to high levels
Solid Lipid Nanoparticle	Relatively low due to rigid structure of solid component	Moderate efficiency	High- solid matrix improves stability	Low toxicity- composed of safe lipids	High- scalability is simple and less complex	Low-moderate
Nanostructured Lipid Carriers	High- lipid matrix allows greater drug encapsulation	High efficiency	High- reduced rigid crystallization allows for minimized drug expulsion	Low toxicity- good biocompatibility	High scalability- but formation is still more complex	Moderate

While SLNs and liposomes all have their unique attributes in delivering medication for Glioblastoma brain tumors, NLCs may offer more balanced overall performance among the three systems. The comparison among the nanoparticles across drug loading capacity, BBB penetration efficiency, stability, toxicity, scalability, and cost is shown (Table 2). Having both solid and liquid lipids within their matrix allows for higher drug loading capacity compared to SLNs, for instance, which have a more rigid structure due to the composition of solid lipids. This means they are more likely to have drug expulsion as a result of not being able to have space for more drug loading. Additionally, NLCs demonstrate improved stability while maintaining lower levels of toxicity, compared to liposomes, which have a

more phospholipid structure and chemically need to be looked after and observed carefully.

However, despite these findings, limited clinical trial data on NLCs prevent the certainty that NLCs are the most effective lipid nanoparticle. Therefore, while NLCs pose improved benefits and are a promising treatment for Glioblastoma brain tumors, further clinical evidence is needed before any definitive conclusions are established.

Conclusion

With high-risk tumors like Glioblastoma, it is crucial to have medicine that will tackle the tumors effectively. While there are thorough options in healthcare to approach Glioblastoma, nanoparticles have been the most promising, specifically lipid nanoparticles. The ability of lipid nanoparticles to cross the BBB effectively, along with efficient drug delivery, has been effective in delivering targeted treatment. SLNs, Liposomes, and NLCs are all effective in their own ways. However, NLCs in particular can have a more flexible crystallized structure, allowing them to have a more targeted treatment and carry higher drug loads. Their biocompatibility and ability to cross the BBB all give them an advantage in treatments. Proceeding with NLCs for Glioblastoma treatment will not only be effective for targeting brain tumors, but also minimize the side effects for patients. Although chemotherapy can slow tumor progression, they often produce healthy tissue and can create significant side effects. Beyond improving outcomes for patients with brain tumors, advances in nanoparticle-based treatments could influence the future of cancer therapy worldwide by creating safer, more precise, and more accessible methods of drug delivery for many difficult-to-treat diseases. Future research can focus more on drug precision and the long-term safety of nanoparticles. Researchers can also focus on combining nanoparticles for immunotherapy or gene therapy to create more selective treatments for patients diagnosed with Glioblastoma.

Acknowledgments

I would like to thank Dr. Kristina Lilova and Professor Virgel Torremocha for their constant guidance throughout the research process. Their remarkable wisdom and patience will never be forgotten.

References

- Qiu, Z.; Yu, Z.; Xu, T.; Wang, L.; Meng, N.; Jin, H.; Xu, B. Novel Nano-Drug Delivery System for Brain Tumor Treatment. *Cells* **2022**, *11* (23), 3761. <https://doi.org/10.3390/cells11233761>.
- American Cancer Society. Chemotherapy Side Effects. <https://www.cancer.org/cancer/managing-cancer/treatment-types/chemotherapy/chemotherapy-side-effects.html>.
- Kumar, L. A.; Pattnaik, G.; Satapathy, B. S.; Swapna, S.; Mohanty, D. Targeting Brain Tumor: Nanocarrier-Based Drug Delivery Platforms, Opportunities, and Challenges. *J. Pharm. Bioallied Sci.* **2021**, *13* (2), 172–177. https://doi.org/10.4103/jpbs.JPBS_239_20.
- Lim, S. H.; Yee, G. T.; Khang, D. Nanoparticle-Based Combinational Strategies for Overcoming the Blood-Brain Barrier and Blood-Tumor Barrier. *Int. J. Nanomedicine* **2024**, *19*, 2529–2552. <https://doi.org/10.2147/IJN.S450853>.

5. Mangraviti, A.; Gullotti, D.; Tyler, B.; Brem, H. Nanobiotechnology-Based Delivery Strategies: New Frontiers in Brain Tumor Targeted Therapies. *J. Controlled Release* **2016**, *240*, 443–453. <https://doi.org/10.1016/j.jconrel.2016.03.031>.
6. Meola, A.; Rao, J.; Chaudhary, N.; Sharma, M.; Chang, S. D. Gold Nanoparticles for Brain Tumor Imaging: A Systematic Review. *Front. Neurol.* **2018**, *9*, 328. <https://doi.org/10.3389/fneur.2018.00328>.
7. UT Southwestern Medical Center. Nanotechnology Helps Chemo Pass the Blood-Brain Barrier. <https://www.utsouthwestern.edu/newsroom/articles/year-2023/oct-nanotechnology-helps-chemo.html>.
8. Mayo Clinic. Glioblastoma - Symptoms and Causes. <https://www.mayoclinic.org/diseases-conditions/glioblastoma/symptoms-causes/syc-20569077>.
9. Cancer Research UK. Glioblastoma. <https://www.cancerresearchuk.org/about-cancer/brain-tumours/types/glioblastoma>
10. Tamimi, A. F.; Juweid, M. Epidemiology and Outcome of Glioblastoma. In *Glioblastoma*; De Vleeschouwer, S., Ed.; Codon Publications: Brisbane, **2017**. <https://www.ncbi.nlm.nih.gov/books/NBK470003/>.
11. Rose, L. Glioma vs. Glioblastoma: What's the Difference? MD Anderson Cancer Center. <https://www.mdanderson.org/cancerwise/glioma-vs--glioblastoma--what-is-the-difference-in-these-brain-tumors-treatment-diagnosis.h00-159537378.html>.
12. University of California Museum of Paleontology. Types of Mutations. Understanding Evolution. <https://evolution.berkeley.edu/dna-and-mutations/types-of-mutations/>.
13. Olafson, L. R.; Gunawardena, M.; Nixdorf, S.; McDonald, K. L.; Rapkins, R. W. The Role of TP53 Gain-of-Function Mutation in Multifocal Glioblastoma. *J. Neuro-Oncol.* **2020**, *147* (1), 37–47. <https://doi.org/10.1007/s11060-019-03318-5>.
14. MedlinePlus. How Do Genes Control the Growth and Division of Cells? MedlinePlus Genetics. <https://medlineplus.gov/genetics/understanding/howgeneswork/genesanddivision/>.
15. Dotiwala, A. K.; McCausland, C.; Samra, N. S. Anatomy, Head and Neck, Blood Brain Barrier. StatPearls Publishing, 2023. <https://www.ncbi.nlm.nih.gov/books/NBK519556/>.
16. Wu, D.; Chen, Q.; Chen, X.; Han, F.; Chen, Z.; Wang, Y. The Blood-Brain Barrier: Structure, Regulation, and Drug Delivery. *Signal Transduction Targeted Ther.* **2023**, *8* (1), 1–27. <https://doi.org/10.1038/s41392-023-01481-w>.
17. Steeg, P. S. The Blood-Tumour Barrier in Cancer Biology and Therapy. *Nat. Rev. Clin. Oncol.* **2021**, *18* (11), 696–714. <https://doi.org/10.1038/s41571-021-00529-6>.
18. Arvanitis, C. D.; Ferraro, G. B.; Jain, R. K. The Blood-Brain Barrier and Blood-Tumour Barrier in Brain Tumours and Metastases. *Nat. Rev. Cancer* **2020**, *20* (1), 26–41. <https://doi.org/10.1038/s41568-019-0205-x>.
19. Gyanani, V.; Haley, J. C.; Goswami, R. Challenges of Current Anticancer Treatment Approaches with Focus on Liposomal Drug Delivery Systems. *Pharmaceuticals* **2020**, *14* (9), 835. <https://doi.org/10.3390/ph14090835>.
20. Mansoori, B.; Mohammadi, A.; Davudian, S.; Shirjang, S.; Baradaran, B. The Different Mechanisms of Cancer Drug Resistance: A Brief Review. *Adv. Pharm. Bull.* **2017**, *7* (3), 339–348. <https://doi.org/10.15171/apb.2017.041>.
21. Zhao, C.; Zhu, X.; Tan, J.; Mei, C.; Cai, X.; Kong, F. Lipid-Based Nanoparticles to Address the Limitations of GBM Therapy by Overcoming the Blood-Brain Barrier, Targeting Glioblastoma Stem Cells, and Counteracting the Immunosuppressive Tumor Microenvironment. *Biomed. Pharmacother.* **2024**, *171*, 116113. <https://doi.org/10.1016/j.biopha.2023.116113>.
22. Ji, A.; Xu, M.; Pan, Y.; Diao, L.; Ma, L.; Qian, L.; Cheng, J.; Liu, M. Lipid Microparticles Show Similar Efficacy with Lipid Nanoparticles in Delivering mRNA and Preventing Cancer. *Pharm. Res.* **2023**, *40* (1), 265–279. <https://doi.org/10.1007/s11095-022-03445-1>.
23. Jnaidi, R.; Almeida, A. J.; Gonçalves, L. M. Solid Lipid Nanoparticles and Nanostructured Lipid Carriers as Smart Drug Delivery Systems in the Treatment of Glioblastoma Multiforme. *Pharmaceutics* **2020**, *12* (9), 860. <https://doi.org/10.3390/pharmaceutics12090860>.
24. Scioli Montoto, S.; Muraca, G.; Ruiz, M. E. Solid Lipid Nanoparticles for Drug Delivery: Pharmacological and Biopharmaceutical Aspects. *Front. Mol. Biosci.* **2020**, *7*, 587997. <https://doi.org/10.3389/fmolb.2020.587997>.
25. Ghasemiyeh, P.; Mohammadi-Samani, S. Solid Lipid Nanoparticles and Nanostructured Lipid Carriers as Novel Drug Delivery Systems: Applications, Advantages, and Disadvantages. *Res. Pharm. Sci.* **2018**, *13* (4), 288–303. <https://doi.org/10.4103/1735-5362.235156>.
26. Grillone, A.; Battaglini, M.; Moscato, S.; Mattii, L.; de Julián Fernández, C.; Scarpellini, A.; Ciofani, G. Nutlin-Loaded Magnetic Solid Lipid Nanoparticles for Targeted Glioblastoma Treatment. *Nanomedicine* **2018**, *14* (6), 727–752. <https://doi.org/10.2217/nnm-2018-0436>.
27. Raju, R.; Abuwatfa, W. H.; Pitt, W. G.; Husseini, G. A. Liposomes for the Treatment of Brain Cancer: A Review. *Pharmaceutics* **2023**, *16* (8), 1056. <https://doi.org/10.3390/ph16081056>.
28. Liu, P.; Chen, G.; Zhang, J. A Review of Liposomes as a Drug Delivery System: Current Status of Approved Products, Regulatory Environments, and Future Perspectives. *Molecules* **2023**, *27* (4), 1372. <https://doi.org/10.3390/molecules27041372>.
29. Lombardo, D.; Kiselev, M. A. Methods of Liposomes Preparation: Formation and Control Factors of Versatile Nanocarriers for Biomedical and Nanomedicine Application. *Pharmaceutics* **2022**, *14* (3), 543. <https://doi.org/10.3390/pharmaceutics14030543>.
30. Xu, M.; Li, G.; Zhang, H.; Chen, X.; Li, Y.; Yao, Q.; Xie, M. Sequential Delivery of Dual Drugs with Nanostructured Lipid Carriers for Improving Synergistic Tumor Treatment Effect. *Drug Delivery* **2020**, *27* (1), 983–995. <https://doi.org/10.1080/10717544.2020.1785581>.
31. Horta, M.; Soares, P.; Sarmiento, B.; Leite Pereira, C.; Lima, R. T. Nanostructured Lipid Carriers for Enhanced Batimastat Delivery Across the Blood-Brain Barrier: An *In Vitro* Study for Glioblastoma Treatment. *Drug Delivery Transl. Res.* **2025**, *15* (8), 2794–2813. <https://doi.org/10.1007/s13346-024-01775-8>.
32. Pudlarz, A.; Szemraj, J. Nanoparticles as Carriers of Proteins, Peptides, and Other Therapeutic Molecules. *Open Life Sci.* **2018**, *13*, 285–298. <https://doi.org/10.1515/biol-2018-0035>.

■ Author

Intrigued by the brain's complexities, Saimanasaa Viswanathan will study neuroscience on a pre-medical track at the University of Texas at Dallas. Outside of school, she plays violin in her school and charity orchestra. She hopes to become a physician while continuing her research and passion for music.

Low Self-Esteem and the School Environment: A Comparative Study of the Risk Factors Among Adolescents

Samarth Kasaraneni, Trayi Yada, Samuel Xiao

Providence Day School, 5800 Sardis Road, Charlotte, NC, 28270, USA; samarth.kasaraneni@gmail.com

ABSTRACT: Adolescence marks some of the most formative and crucial years of an individual's life. Consequently, during this period, many adolescents are prone to changes in their social and emotional well-being. With a majority of adolescents spending a considerable portion of their time in school, their academic setting plays a pivotal role in shaping their identity. During this period, low self-esteem is a common mental health issue that often goes unspoken about. The purpose of this study is to address common risk factors that affect an adolescent's self-esteem and compare them across school settings. The risk factors were divided into two groups: academic-based risk factors and social-based risk factors. To accomplish the goal of this study, a survey link was made public and shared with students in schools (public and independent) across five states (North Carolina, South Carolina, Ohio, Illinois, and Arizona). In total, we collected responses from exactly 400 students by placing a cap on the responses after we reached a certain threshold. Adolescents attending independent schools tended to be more inclined towards the academically based risk factors, whereas adolescents attending public schools were inclined towards the social-based risk factors.

KEYWORDS: Social and Behavioral Science, Educational Psychology, School Environment, Adolescent Mental Health, Self-Esteem.

■ Introduction

Adolescence, typically occurring in individuals between ages 10 and 19, is defined as a rapid period of physical growth and intellectual development.¹ Consequently, high school years serve as a crucial, formative time for young teenagers to mature into adolescents. Beyond simply defining the identity of the individual, however, adolescence also marks the crucial years during which teenagers begin developing their mental health and emotional stability.² During this period, self-esteem, commonly defined as the totality of an individual's thoughts and feelings regarding themselves, is paramount in either corroborating or hindering the emotional development of the young adolescent.³ From a holistic view, self-esteem has the potential to influence aspects of adolescents' lives beyond just their emotions, encompassing both academic performance and social comfort.⁴ Students with higher self-esteem are often better equipped to cope with stress, build healthy relationships, and maintain emotional resilience in times of hardship. On the other hand, low self-esteem can be characterized as a lack of confidence in oneself and difficulty recognizing one's own worth.⁵ Low self-esteem, if left unaddressed, can develop into negative thoughts and potentially harmful ideations for the maturing adolescent.⁶ With a majority of young adolescents spending a considerable portion of their time in school, the academic environment that they are exposed to has a significant role in determining their emotional and physical development, as the effects of daily interactions accumulate over time and shape the individuality and beliefs of the adolescent.⁷

However, when adolescents lack adequate support or face persistent academic or social pressure, they can be subject to diminished self-esteem and fragile mental health.⁸ As a result, the type of school an adolescent attends (public or indepen-

dent) can have a considerable impact on their future well-being and overall self-esteem.⁹ Whether through differences in resources, access to professionals, class sizes, teaching styles, or social environments, public schools and independent schools tend to vary in ways that may influence adolescent self-esteem. Consequently, it is the students who face the brunt of these differences. These differing school environments can introduce both risk and protective factors that may influence students' self-esteem by shaping their academic pressure, sense of social belonging, and access to mental health support. Protective factors are conditions that may decrease the likelihood of a negative outcome or act as a buffer to the symptoms of the condition.¹⁰ Risk factors, on the other hand, are conditions that may increase the likelihood of a negative outcome: in this case, low self-esteem.¹⁰ Schools, irrespective of being public or independent, typically offer at least one type of mental health service to their students.¹¹ Because we have attended both public and private schools and have formed close friendships with peers in both school settings, we developed an interest in examining the underlying differences between the school environments and their corresponding effects on students' mental health. Whether it is through offering mentorship or having spaces for social support, most schools provide basic safety nets for students. Even so, the unspoken reality is that many high school adolescents, on average, quietly suffer from low self-esteem due to a variety of factors.¹² That is the purpose of this study: to analyze how risk factors for low self-esteem vary between students attending public and independent schools.

Although studies have already been done regarding overall teen mental health in school, there have been a very limited number of comparative studies to date examining the risk factors for low self-esteem between students attending public vs.

independent schools. We hypothesize that students attending either type of school will experience similar proportions of academically related risk factors for low self-esteem, whereas students attending public schools will experience higher proportions of social-related risk factors for low self-esteem than their peers in independent schools. The results of this research can help parents, educators, and policymakers better understand the mental health aspect of school and its potential underlying effects on students.

■ **Methods**

To accomplish the goal of this study, a survey link was made public and shared with students in schools (public and independent) across five states (North Carolina, South Carolina, Ohio, Illinois, and Arizona). We opted to craft our own survey instead of proceeding with a standardized survey, as we could build on top of our prior research analyzing the most important risk factors of low self-esteem for adolescents. Additionally, through this approach, we could delve into a deep comparative analysis of these specific risk factors. In the survey, the following questions were asked:

- I. Grade level
 - A. 9th
 - B. 10th
 - C. 11th
 - D. 12th
- II. Which type of school do you attend?
 - A. Public
 - B. Independent
- III. Which US state are you from?
 - A. _____
- IV. Low self-esteem can be characterized as a lack of confidence in oneself and difficulty recognizing one’s own worth. Have you experienced a period of low self-esteem (spanning anytime between multiple days to over several years) in your life?
 - A. Yes
 - B. No
- V. If you answered “Yes” to the previous question, please select all of the risk factors that you believe have been associated with your low self-esteem. If you answered “No” to the previous question, you do not need to respond to this.
 - A. Perfectionism / High Expectations
 - B. Chronic Stress
 - C. Harsh Criticism
 - D. Social Rejection
 - E. Unstable Home Environment
 - F. Social Media

In total, we collected responses from exactly 400 students by placing a cap on the responses after we reached a certain threshold. Specifically, we got 50 responses from each of the following subgroups: 9th graders in public school, 10th graders in public school, 11th graders in public school, 12th graders in public school, 9th graders in an independent school, 10th graders in an independent school, 11th graders in an independent school, 12th graders in an independent school. To spread the study, we initially shared our research survey with a subset

of students and later relied on individuals spreading the survey to their peers. No personal information, such as the respondents’ names, addresses, or emails, was collected to preserve confidentiality. Once we reached 50 responses from each of these subgroups, we stopped collecting more responses to keep the sample size even across all of the subgroups.

■ **Results and Discussion**

From the sample of 400 students surveyed, 389 (97.25%) students responded by stating that they have currently/recently experienced lowered self-esteem during a given period of time.

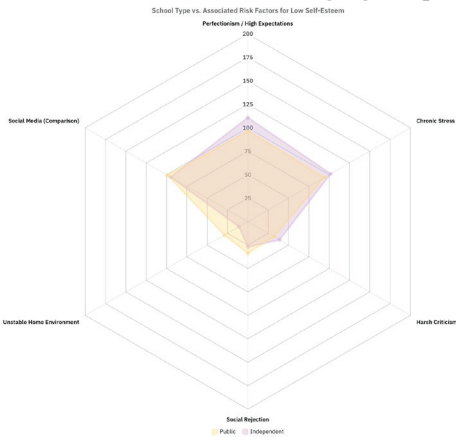


Figure 1: Radar plot showing the distribution of selected risk factors among the 389 students who reported having experienced a period of low self-esteem. This is intended to highlight the differences between potential stressors across the two different school environments. The yellow data points visible on the axes refer to the frequency of public-school students selecting each risk factor. Likewise, the purple data points visible on the axes refer to the frequency of independent-school students selecting each risk factor.

Among those 389 survey respondents who reported having experienced a period of low self-esteem, their responses regarding perceived risk factors associated with lowered self-esteem are presented in Figure 1. The respondents were grouped in this graph based on their associated school type. This general overview highlights the trend that among high school adolescents attending independent schools, the students tended to express higher academically related risk factors, including perfectionism, chronic stress, and harsh criticism, as evidenced by the purple data points being located farther away from the center than the yellow points for these factors. On the other hand, adolescents attending public schools were predisposed to more socially based stressors for low self-esteem, such as social rejection, an unstable home environment, and constant comparison via social media use, as indicated by the yellow data points being farther away from the center than the purple data points for these axes.

Table 1A: List of all the frequencies for each selected risk factor in a public school setting. The number of students who reported having experienced a period of low self-esteem from each subgroup is placed in parentheses (e.g., 9th Grade (50)). It is also important to note that all of the values in a given row may not add up to the value in the parentheses, as students were allowed to select multiple risk factors and not just one.

Public	Perfectionism / High Expectations	Chronic Stress	Harsh Criticism	Social Rejection	Unstable Home Environment	Social Media (Comparison)
9th Grade (50)	18	14	6	11	7	20
10th Grade (48)	23	22	6	10	8	27
11th Grade (47)	26	25	11	7	7	26
12th Grade (48)	30	33	9	6	7	26

Table 1B: List of all the frequencies for each selected risk factor in an independent school setting. The number of students who reported having experienced a period of low self-esteem from each subgroup is placed in parentheses (e.g., 10th Grade (47)). It is also important to note that all of the values in a given row may not add up to the value in the parentheses, as students were allowed to select multiple risk factors and not just one.

Independent	Perfectionism / High Expectations	Chronic Stress	Harsh Criticism	Social Rejection	Unstable Home Environment	Social Media (Comparison)
9th Grade (50)	22	14	9	10	3	20
10th Grade (47)	26	22	8	7	1	24
11th Grade (50)	31	32	12	6	3	26
12th Grade (49)	32	34	10	4	4	25

Tables 1A and 1B show a more detailed breakdown of each of the risk factors by grade level for a given school type. This data served as the basis for the radar plot (Figure 1) as well as for the relative frequency graphs (Figures 2, 3, 4, and 5).

Comparative Analysis Across School Types:

In order to compare the difference in perceived risk factors across school types, chi-square analyses were utilized.

Table 2: Analysis of each of the risk factors across school types. χ^2 signifies how much an observed value differs from an expected value.¹³ The p-value is used to deduce whether or not a result is statistically significant. $p < 0.05$ indicates that a result is statistically significant; $p > 0.05$ indicates that a result is not statistically significant. As indicated by the table, the only risk factor that is statistically significant is living in an unstable home environment.

	Chi-Square Analysis	
	χ^2	p-value
Perfectionism	1.34	0.247
Chronic Stress	0.31	0.578
Harsh Criticism	0.51	0.474
Social Rejection	0.81	0.367
Unstable Home Environment	9.35	0.002
Social Media	0.21	0.648

As shown in Table 2, the p-scores revealed non-statistically significant results across five of the six risk factors. Even so, a statistically significant difference was derived from the risk factor of an unstable home environment ($p = 0.002$); thus, it can be concluded that students who listed living in an unstable home environment as a potential stressor were significantly more likely to come from a public school than an independent school. Furthermore, although it is difficult to affirmatively state that students from a certain school type are inclined towards the risk factors of perfectionism, chronic stress, harsh criticism, social rejection, or social media use, coming from an unstable home environment poses an unequal risk for students attending public schools.

To continue, examining the risk factors most favored across both school types, perfectionism stands out as the most frequently selected, with 53.5% of the 389 students reporting it as at least one potential stressor. Following perfectionism, chronic stress and comparison via social media use were the next most frequently selected stressors, with each risk factor garnering 50.4% and 49.9% selection, respectively. After these three risk factors, a sharp decline becomes visible in the selection rate for the last three. Harsh criticism follows as the next most selected risk factor, with 18.3% of adolescents selecting it. Lastly, social rejection and living in an unstable home were reported by 15.7% and 10.3% of the high school adolescents.

Comparative Analysis Across Grade Levels:

As previously stated, although chi-square analyses did not provide statistically significant results across all potential risk

factors, visible trends, as shown in Figures 2, 3, 4, and 5, can still be observed across grade levels and school types. Firstly, as the grade level of students progressed higher (9th through 12th), adolescents, irrespective of which type of school they attended, had a stronger inclination to associate perfectionism and chronic stress as potential stressors for low self-esteem. To exemplify this, 36% of the 9th graders attending public school who reported having experienced a period of low self-esteem in their lives selected perfectionism as a stressor for low self-esteem. Upon observation of 10th, 11th, and 12th graders attending public schools, 48%, 55%, and 63% of students, respectively, selected perfectionism as a potential stressor for low self-esteem. This same increasing pattern can be found among students attending independent schools reporting perfectionism as a potential risk factor for low self-esteem, as well as for students of both school types reporting chronic stress.

In general, across all four grade levels, the risk factors of perfectionism, chronic stress, and comparison via social media use were consistently selected with the highest proportions as potential stressors for low self-esteem. When comparing the relative frequency bars in Figure 2 to those of the other figures, the overall proportion of high school freshmen selecting potential risk factors for low self-esteem is lower compared to students in other grades. Furthermore, when comparing the proportion of students who selected academic-based stressors as potential risk factors for low self-esteem, juniors (Figure 4), regardless of school type, reported higher proportions than freshmen (Figure 2) and sophomores (Figure 3). Specifically, the proportion of students selecting perfectionism, chronic stress, and harsh criticism as potential risk factors for low self-esteem gradually increased in students through grade levels. The one exception to this trend can be found for the risk factor of harsh criticism among independent school freshmen and sophomores, for whom proportions of 0.18 and 0.17 were selected, respectively.

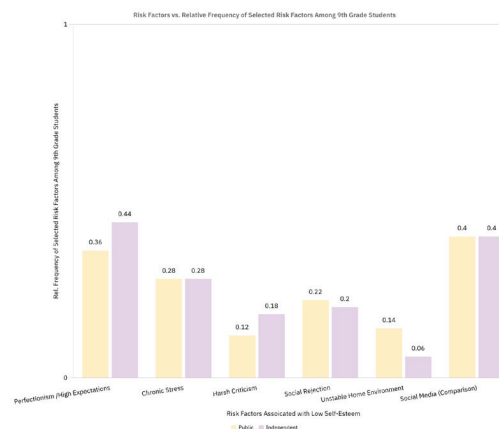


Figure 2: Relative frequency graph for each of the risk factors affecting 9th graders in public and independent school settings. Each of the bars is relative to the number of 9th graders in the associated school setting who reported having a period of low self-esteem in their lives. The yellow bars represent the relative frequencies for public school students, and the purple bars represent the relative frequencies for independent school students.

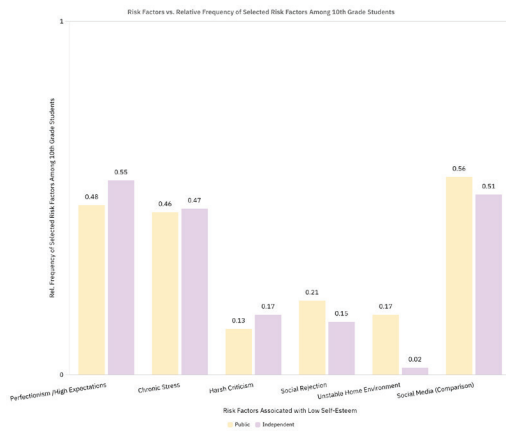


Figure 3: Relative frequency graph for each of the risk factors affecting 10th graders in public and independent school settings. Each of the bars is relative to the number of 10th graders in the associated school setting who reported having a period of low self-esteem in their lives. The yellow bars represent the relative frequencies for public school students, and the purple bars represent the relative frequencies for independent school students.

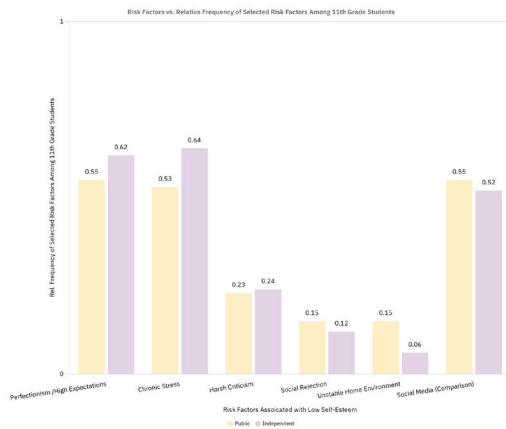


Figure 4: Relative frequency graph for each of the risk factors affecting 11th graders in public and independent school settings. Each of the bars is relative to the number of 11th graders in the associated school setting who reported having a period of low self-esteem in their lives. The yellow bars represent the relative frequencies for public school students, and the purple bars represent the relative frequencies for independent school students.

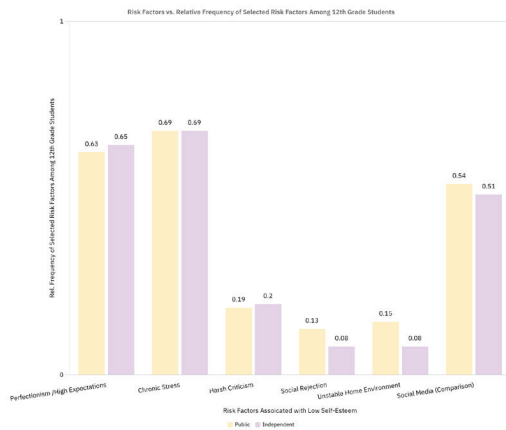


Figure 5: Relative frequency graph for each of the risk factors affecting 12th graders in public and independent school settings. Each of the bars is relative to the number of 12th graders in the associated school setting who reported having a period of low self-esteem in their lives. The yellow bars represent the relative frequencies for public school students, and the purple bars represent the relative frequencies for independent school students.

Two visible differences between the relative frequency graphs of selected risk factors across juniors and seniors are that seniors (Figure 5) reported higher proportions of perfectionism and chronic stress than juniors (Figure 4). Conversely, seniors reported lowered proportions of harsh criticism, social rejection, and comparison via social media. Furthermore, similar to the overall pattern of higher proportions of independent school students associating themselves with academic-based risk factors and higher proportions of public school students associating themselves with social-based risk factors, the same trend follows with the seniors, as evidenced in Figure 5. The sole exception is that equal proportions of seniors across school types selected chronic stress as a potential risk factor for low self-esteem.

Additionally, with regard to social media use as a form of comparison, the percentage of high school adolescents in 10th, 11th, and 12th grade reporting it as a potential stressor for low self-esteem varied between 51% and 56%, irrespective of the school type.

Lastly, another valuable observation that can be deduced from the given data is that academically related risk factors (including perfectionism, chronic stress, and harsh criticism) were selected at a higher rate than social-based stressors (including social rejection, living in an unstable home environment, and comparison via social media use). Specifically, given the proportions of students selecting each factor, the three academically oriented risk factors averaged to 40.7%, whereas the socially oriented risk factors averaged to 25.3%, signifying that students more frequently reported academically induced stressors as potential risk factors for low self-esteem than socially based ones.

Conclusion

In summary, considering the importance of adolescence on individuals and their overall mental health, high school years mark some of the most formative and pivotal years of a student's life. Consequently, if provided with adequate care and attention, adolescents will be much more likely to develop long-term emotional stability and heightened self-esteem. Fortunately, a majority of schools offer some form of mental health support through either counseling or having social support spaces. However, what often goes unnoticed in academic settings are the hidden battles that many students fight, whether it is against themselves or external pressures. School presents young adolescents with challenges that many have to learn to overcome. As revealed through the data in this research, for example, a large portion of students confront lowered self-esteem levels, with school environments playing a considerable role in these outcomes.

There are a few major trends that can be utilized to help shape potential measures in school environments. To begin, recognizing that academic pressures such as perfectionism and chronic stress increase throughout an adolescent's journey to high school is valuable. This understanding helps highlight the need for early support interventions focused on stress management and potential after-school teacher help sessions. Furthermore, realizing that social media could serve as a plat-

form for persistent comparison against others can be crucial in identifying potential lifestyle changes for an adolescent. The consistent use of social media and its deleterious side effects on self-esteem, irrespective of school type, demonstrates the constant and persistent impact of social media exposure on teens and developing adolescents. These findings align with pre-existing literature stating that adolescent self-esteem is strongly shaped by the stressors that they are most frequently exposed to, such as academic pressures and social comparison. By implementing soft regulations against social media use, adolescents can be trained to limit constant comparison and shift their focus to themselves rather than to others. Lastly, with regard to the risk factor of residing in an unstable home environment, this is a scenario that schools and their administrators should be more vigilant about. While they have no direct power in being able to influence an adolescent's current household situation, they can underscore the implementation of support programs at school that offer emotional guidance to adolescents facing these scenarios. More broadly, this study contributes to the growing body of research supporting that varying academic and social settings across school types contribute differently to an adolescent's emotional development.

In short, a school's impact on adolescents is undeniable: it prepares students for their futures and plays a critical role in the development of an individual's emotional and social well-being.¹⁴ While this study revealed several meaningful trends regarding adolescent self-esteem across school settings, future research involving broader populations and the establishment of causation across different risk factors could further strengthen understanding of how school environments influence student mental health. Despite the obstacles and challenges that a maturing adolescent may face, by recognizing the differing patterns of academically and socially related stressors between public and independent school students, counselors and caregivers can be more vigilant and work to connect with the individual to support them through the journey of adolescence while standing as pillars of support along the way.

■ Acknowledgments

We would like to express our deepest gratitude to all of the students who were willing to participate in our survey. We would also like to extend our heartfelt thanks to all of the teachers and administrators who work to make school a safe and engaging learning environment.

■ References

1. WHO. *Adolescent Health*. World Health Organization. https://www.who.int/health-topics/adolescent-health#tab=tab_1.
2. World Health Organization. *Mental health of adolescents*. World Health Organization. <https://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health>.
3. Department of Sociology. *Self-Esteem: What Is It?* | *SOCY1 Sociology Department I University of Maryland*. Umd.edu. <https://socy.umd.edu/about-us/self-esteem-what-it>.
4. Camel, D. *Improving Students' Academic Performance from the Inside Out: The Relationship between Self-Esteem and Academic Performance. A Research Snapshot*, 2020. https://www.jsums.edu/murc/files/2020/07/Self-EsteemAcademicPerformance.MURC_.Snapshot_June30.2020.pdf.
5. Cherry, K. *11 Signs of Low Self-Esteem*. Verywell Mind. <https://www.verywellmind.com/signs-of-low-self-esteem-5185978>.
6. OVERHOLSER, J. C.; ADAMS, D. M.; LEHNERT, K. L.; BRINKMAN, D. C. Self-Esteem Deficits and Suicidal Tendencies among Adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry* 1995, 34 (7), 919–928. <https://doi.org/10.1097/00004583-199507000-00016>.
7. Barney, L. *The Connection Between Self-Esteem and Academic Success (And How We Boost Both)* – McLean School. McLean School. <https://www.mcleanschool.org/the-connection-between-self-esteem-and-academic-success-and-how-we-boost-both/>.
8. Hensley, S. E. *Study finds school choice improves mental health*. Oklahoma Council of Public Affairs. <https://ocpathink.org/post/independent-journalism/study-finds-school-choice-improves-mental-health>.
9. Chen, Y.; Hinton, C.; VanderWeele, T. J. School Types in Adolescence and Subsequent Health and Well-Being in Young Adulthood: An Outcome-Wide Analysis. *PLOS ONE* 2021, 16 (11), e0258723. <https://doi.org/10.1371/journal.pone.0258723>.
10. CDC. *Risk and Protective Factors*. Adverse Childhood Experiences (ACEs). <https://www.cdc.gov/aces/risk-factors/index.html>.
11. kffnirmitap. *The Landscape of School-Based Mental Health Services* | KFF. KFF. <https://www.kff.org/mental-health/the-landscape-of-school-based-mental-health-services/>.
12. Beresin, E. *Low Self-Esteem in Adolescents: What are the Root Causes?* | *Psychology Today*. Wwww.psychologytoday.com; Psychology Today. <https://www.psychologytoday.com/us/blog/inside-out-outside-in/202206/low-self-esteem-in-adolescents-what-are-the-root-causes>.
13. Hayes, A. *How the Chi-Square Statistic Works*. Investopedia. <https://www.investopedia.com/terms/c/chi-square-statistic.asp>.
14. Mountain Heights Academy. *Mountain Heights Academy | Utah Online Public School Grade 7-12 – How School Affects Mental Health: Understanding the Impact on Students*. Mountainheightsacademy.org. <https://mountainheightsacademy.org/how-school-affects-mental-health-understanding-the-impact-on-students>.

■ Authors

Samarth Kasaraneni is currently a senior at Providence Day School in North Carolina. He hopes to pursue a degree in neuroscience and enter the medical field as a neurosurgeon, with a particular interest in addressing healthcare disparities within underserved communities.

Sam Xiao is currently a senior at Providence Day School, NC. He hopes to pursue a degree in Biomedical Engineering and enter the medical field as an oral surgeon.

Trayi Yada is currently a senior at Providence Day School in North Carolina. She aspires to earn a combined JD/MD degree and pursue a career as a medical lawyer, aiming to work at the intersection of healthcare, law, and advocacy.

Beyond the Pale: How Cognitive Biases Influence Our Preference for White, Tree-based Paper

Seoha Han

North London Collegiate School Jeju, 33 Globaleduro 145 Beongil, Jeju-do, 63644, South Korea; stellabliss09@gmail.com

ABSTRACT: The overwhelming use of bleached, tree-based white paper—a significant contributor to deforestation and toxic pollution—remains unquestioned despite the availability of environmentally superior alternatives like beige, tree-free sugarcane paper. This paper addresses an overlooked, pervasive barrier to eco-friendly consumption, the irrational aversion to tree-free paper based on a non-white color rather than function. This study investigates the extent to which cognitive biases may influence the preference for tree-based, white paper over beige, tree-free sugarcane paper. A Likert-scale survey ($N = 122$) was administered to secondary school students at a school in South Korea, which has adopted sugarcane paper by participating in my environmental initiative. An ANOVA demonstrated significant differences in agreement ratings across cognitive bias statements ($F(10, 1210) = 7.85, p < .001$). The results showed that *Familiarity Bias*, *Aesthetic Bias*, and *Appeal to Popularity* were the biases most strongly associated with white paper preference. Ultimately, this research serves as an initial investigation into the widespread preference for white, tree-based paper. It highlights the potential fruitfulness of addressing psychological factors preventing the adoption of sustainable alternatives, especially for adolescents, who will be the next leaders in climate change mitigation.

KEYWORDS: Psychology, Environmental Psychology, Cognitive Bias, Sustainability, Sugarcane Paper.

■ Introduction

In educational, professional, and personal settings, the use of tree-based, white paper to print worksheets, examinations, official documents, and other materials remains a widely accepted and largely unquestioned practice, quantified by the average office worker's consumption of 10,000 sheets of copy paper per year.¹ However, this widespread preference for white wood-pulp paper directly contributes to massive levels of deforestation each year, as an estimated 53.33% of trees worldwide are logged for paper production.² Since trees absorb significant amounts of carbon, deforestation substantially contributes to a gradual increase in the atmospheric concentration of greenhouse gases and hence to global warming.³ Moreover, the authentic hue of wood pulp is beige; the pure white of tree-based paper is the result of artificial bleaching of wood pulp in highly toxic, non-biodegradable chemicals detrimental to the environment and human health.⁴ An alternative to the standard white wood-pulp paper is sugarcane paper, which is made from bagasse – waste from harvesting sugarcane.⁵ While it is possible to bleach sugarcane paper, it is predominantly sold unbleached, as it is functional in its natural beige hue, and the companies that produce sugarcane paper tend to be environmentally conscious.⁶ Both of these factors make sugarcane paper a more environmentally friendly option.⁵ At my school, which has adopted sugarcane paper in some of its printers by participating in my environmental initiative, I have observed an aversion to sugarcane paper.⁷ For example, a few teachers, upon learning that the printer was filled with sugarcane paper, walked an entire flight of stairs to use the printer on the next floor, which had white paper. Some students cited concerns that they did not like the paper because it was “too beige.” The

school had introduced sugarcane paper during the previous academic year through pilot trials in selected printers, school assemblies, and classroom use. Empirical studies in behavioral economics and psychology have demonstrated that when making decisions related to climate change, individuals often deviate from rational decision-making in consistent, systematic, and predictable ways.⁸ Many of these deviations can be labelled cognitive biases: “systematic tendencies or distortions in human judgement and decision-making.”⁹

In terms of the two modes of cognitive processing proposed by Kahneman, cognitive biases—the mental shortcuts of System 1 thinking—often override the more rational, deliberate System 2 thinking.¹ For example, Dhami *et al.* conducted laboratory experiments where participants made decisions about contributing to a green fund that would reduce the future likelihood of a climate catastrophe.² The study found that present bias, the tendency to overvalue immediate outcomes and undervalue future ones, and loss aversion, where the pain of a loss is felt more intensely than the pleasure of an equivalent gain, decreased the participants' contributions to the green fund.^{3,4}

However, to my knowledge, there have been no papers specifically investigating the reasons behind the societal preference for white tree-based paper despite its environmental cost and the availability of myriad sustainable alternatives, like beige sugarcane paper. To explore this inquiry, I investigated cognitive biases present in students at a school that has adopted the use of sugarcane paper in a few of its printers.

I hypothesize that one or more cognitive biases may significantly contribute to this phenomenon, with the status quo bias—the propensity to choose to maintain the status quo rather than make changes—contributing the most.⁵ This is because

this cognitive bias has been observed in other areas of climate action, as environmentally friendly choices often require active effort to depart from the status quo, such as opting into a green electricity tariff or specifically requesting a vegetarian meal, making them less likely to be adopted.^{6,7} In the case of paper, the status quo bias may hinder the wider adoption of sustainable alternatives like sugarcane paper, as white wood-pulp paper is the status quo, commonly accepted without question.

This investigation aims to explore how insights from cognitive psychology may help identify psychological barriers to the adoption of sustainable alternatives such as sugarcane paper. By understanding such psychological barriers to sustainability, we can develop targeted solutions for individuals and institutions and thus facilitate the widespread adoption of sustainable practices.

■ Methods

Data Collection:

Data were collected through a 5-point Likert-scale survey, titled “Paper Preferences and Perceptions,” and created using Google Forms.⁸ The target population was the students at North London Collegiate School Jeju (age 11 to 18), as they had prior exposure to beige, wood-pulp-free sugarcane paper following its adoption by the school. This study used opportunity volunteer sampling.

A waiver of individual parental consent was implemented, given the research context. The study posed minimal risk since it addressed a non-sensitive topic: preferences for school-provided paper. Moreover, the study—conducted in an educational setting—proceeded with explicit permission of Heads of Houses and teaching staff. Also, students were informed that the survey was not mandatory, which protected student assent, and anonymity was guaranteed.

Upon request, teachers allotted class time for optional survey completion in multiple classes across different year groups they taught, and Heads of Houses provided form time for all year groups to participate. This ensured that students across all year groups had access to the survey. Out of an estimated total population of 1400, 122 responded, which represents approximately 9% of the population. The low response rate was partially due to the survey being accessible to only certain classes and houses, as some teachers did not respond to emails requesting to share the survey. Despite the low response rate, the collected sample size is statistically sufficient to provide meaningful analysis, yielding a 8.5% Margin of Error with a 95% confidence level for a population of 1400. This is considered sufficient for exploratory research.⁹

The survey consisted of 2 sections. In the first section, students rated their agreement with 11 statements, each designed to measure a specific cognitive bias, on a Likert scale from 1 (strongly disagree) to 5 (strongly agree). The second section contained 7 statements assessing students’ perceptions of change, information, and values, also rated on the same 5-point scale. Single-item measures were used for each cognitive bias to maintain survey brevity and minimize participant fatigue within the context of a voluntary school-based survey.

The survey was designed with neutral titles and descriptions so that respondents were unaware that the study was investigating cognitive biases. For instance, the survey, its first section, and its second section were respectively titled “Paper Preferences and Perceptions,” “Factors Influencing Choice of Paper,” and “Perceptions of Change, Information, and Values”. The description contained a brief, factual explanation of sugarcane paper, devoid of persuasive or evaluative language. This approach was intended to prevent participants from consciously adjusting their responses to align with socially desirable or ‘unbiased’ thinking.

Data Handling:

Google Sheets (where the Google Form responses were automatically recorded) and RStudio (Appendix A) were used for data processing. Among the total of 125 responses to each statement, it was observed that three participants selected “3” (neutral) for all 18 statements. As it is highly unlikely that these respondents were neutral towards all statements in the form, their responses were removed. Consequently, 122 responses were included in the statistical analysis.

The Google Sheet was imported into RStudio and divided into two tables: “results_part1” for all the participant ratings of the 11 cognitive bias statements, and “results_part2” for all the participant ratings of the remaining 7 statements regarding change, information, and values. The mean values for the 11 cognitive bias statements and the remaining 7 statements were subsequently calculated.

Statistical Analysis:

To determine whether mean agreement ratings differed significantly across the cognitive bias statements investigated, a within-subjects Analysis of Variance (ANOVA) test was conducted. Subsequently, a pairwise t-test was performed with a Holm correction. Each test was interpreted using the standard $p < .05$ criterion for significance.

■ Results and Discussion

Results:

Means and Standard Deviations for the 11 cognitive bias statements and the remaining 8 statements are presented in Appendix C. On average, participants agreed with Familiarity Bias ($M = 3.6$, $SD = 1.1$) the most, and Authority Bias ($M = 2.8$, $SD = 1.2$) the least.

The ANOVA test showed significant differences in mean agreement ratings across cognitive bias statements, with a modest effect size, $F(10, 1210) = 7.85$, $p < .001$, $\eta^2 = .06$.

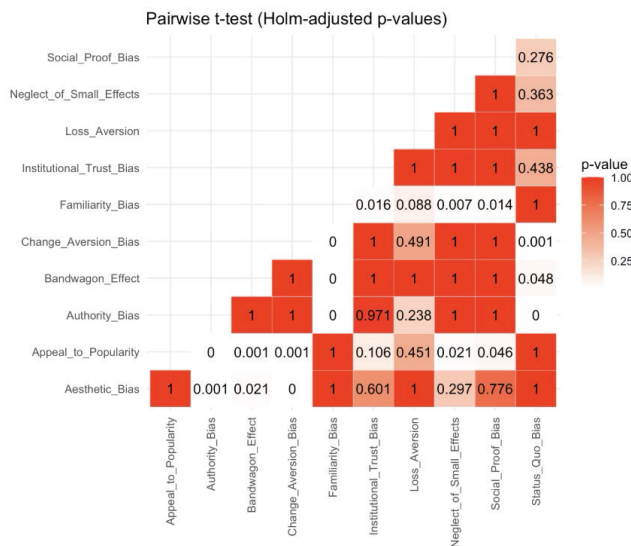


Figure 1: Heatmap of Holm-Adjusted P-Values from Pairwise T-Tests of Cognitive Biases. The heatmap indicates the statistical significance of pairwise differences between biases after Holm correction. Familiarity Bias, Appeal to Popularity, and Aesthetic Bias received relatively higher ratings, while Authority Bias, Bandwagon Effect, and Change Aversion Bias were rated lower overall.

The post hoc results in Figure 1 revealed several significant pairwise mean differences between the biases. *Familiarity Bias* was rated significantly higher than *Social Proof Bias* ($p = .003$), *Neglect of Small Effects* ($p = .005$), and *Institutional Trust Bias* ($p = .012$). *Appeal to Popularity* was rated significantly higher than *Social Proof Bias* ($p = .010$), *Neglect of Small Effects* ($p = .015$), *Institutional Trust* ($p = .035$), *Change Aversion* ($p < .001$), *Bandwagon Effect* ($p = .005$), and *Authority Bias* ($p < .001$). *Aesthetic Bias* was rated significantly higher than both *Change Aversion Bias* ($p = .002$) and *Authority Bias* ($p < .001$).

In contrast, *Change Aversion Bias* was rated significantly lower than *Status Quo Bias* ($p = .001$) and *Familiarity Bias* ($p < .001$). The *Bandwagon Effect* was rated significantly lower than *Familiarity Bias* ($p = .001$). *Authority Bias* was rated significantly lower than both *Status Quo Bias* ($p < .001$) and *Familiarity Bias* ($p < .001$). Overall, these findings indicate that *Familiarity Bias*, *Appeal to Popularity*, and *Aesthetic Bias* had generally high mean agreement ratings, while *Authority Bias*, *Bandwagon Effect*, and *Change Aversion Bias* were rated lower on average. Therefore, the null hypothesis—that there would be no statistically significant difference in the mean ratings of the cognitive biases—is rejected. The alternative hypothesis is thus accepted, which suggests that several cognitive biases may be associated with students' preference for white paper. These findings therefore suggest that certain cognitive biases may be more strongly associated than others with adolescents' preference for white, tree-based paper.

Discussion:

The findings support the hypothesis that students' preference for white, tree-based paper over beige sugarcane paper may be associated with several of the cognitive biases investigated. However, it does not substantiate the part of the hypothesis that status quo bias most significantly contributes

to the explored phenomenon, as participants showed the highest mean ratings for *Familiarity Bias* and *Appeal to Popularity*. This could be explained in light of the adolescent age range of the sample. According to Pinho *et al.*, adolescents' personal norms are significantly shaped by majority norms. Hence, one possible interpretation is that participants may have gravitated towards what is familiar and popular—such as the widespread use of white paper—to fit in.¹¹

Similarly, the high rating for *Aesthetic Bias* could be interpreted in the context of adolescents' heightened susceptibility to self-consciousness.¹² Adolescents are often particularly attentive to physical appearance and self-presentation. This may have extended to the materials they use, potentially contributing to their preference for pristine white paper over the beige alternative, although this interpretation was not directly tested in the study.

In contrast, the results demonstrated that *Authority Bias* had the lowest mean agreement, which may reflect patterns identified in previous research on adolescence. As they seek greater independence, adolescents tend to develop a propensity to question and defy authority.¹³ Consequently, the endorsements of sugarcane paper by authority (expressed as “a respected institution” by the survey) may have had little impact on students' paper preferences.

There was a surprisingly significant difference in mean ratings between the ostensibly similar biases, *Familiarity Bias* and *Change Aversion Bias*. While *Familiarity Bias* had the highest mean agreement, *Change Aversion Bias* had the second lowest. Self-serving bias may partially explain this discrepancy, although this interpretation was not directly tested in the study: participants may have perceived the *Change Aversion Bias* statement “I do not feel comfortable with the idea of switching because I don't like change” as admitting a character flaw. As people tend to avoid admitting failure, participants may have been unwilling to agree with the *Change Aversion Bias* statement.¹⁴ On the other hand, the *Familiarity Bias* statement, “I feel uncertain about sugarcane paper because I'm more familiar with white paper”, was perhaps less likely to be interpreted as a personal shortcoming.

Likewise, while the statements of the *Bandwagon Effect*, *Authority Bias*, and *Social Proof Bias* were phrased in a way that implied a lack of independent thought—a potential character flaw—the statement for *Appeal to Popularity* was more generic and less direct. Thus, students—adolescents in their pursuit of autonomy—may have been more inclined to reject the statements for the *Bandwagon Effect*, *Authority Bias*, and *Social Proof Bias* statements and agree with the *Appeal to Popularity* statement, despite the potential similarity of the four cognitive biases, although this interpretation remains speculative.

These findings align with the mean ratings of other statements investigated in the study, which provided further context for the participants' attitudes towards sugarcane paper. For example, the high mean ratings (3.7 for both) for “I do not mind trying new materials if they are more sustainable” and “Sustainability is one thing I consider when choosing materials, but not the only thing” may indicate generally positive attitudes toward sustainability. However, the comparatively lower

mean rating (3.2) for "I would support my school/workplace adopting sugarcane paper, even if it felt unfamiliar at first" may support the possibility that *Familiarity Bias* is associated with students' white paper preference.

Table 1: Means and Standard Deviations for Each Cognitive Bias. Familiarity Bias, Appeal to Popularity, and Aesthetic Bias received the highest mean ratings, while Authority Bias and Change Aversion Bias received the lowest.

Cognitive bias	M (1 d.p.)	SD (1 d.p.)
Status Quo Bias	3.4	1.2
Loss Aversion	3.2	1.2
Aesthetic Bias	3.4	1.3
Bandwagon Effect	3.0	1.2
Appeal to Popularity	3.5	1.2
Familiarity Bias	3.6	1.1
Authority Bias	2.8	1.2
Neglect of Small Effects	3.0	1.3
Institutional Trust Bias	3.1	1.2
Social Proof Bias	3.0	1.1
Change Aversion Bias	2.9	1.2

Table 2: Means and Standard Deviations for Each Statement About Change, Information, and Values. Participants generally agreed with statements supporting sustainability, but showed slightly less agreement with adopting sugarcane paper if it felt unfamiliar.

Statement	M (1 d.p.)	SD (1 d.p.)
I do not mind trying new materials if they are more sustainable.	3.7	1.1
I would support my school/workplace adopting sugarcane paper, even if it felt unfamiliar at first.	3.2	1.2
I feel confident I understand the differences between sugarcane paper and white paper.	3.5	1.1
I need more information before deciding whether sugarcane paper is a good choice.	3.4	1.2
I see myself as someone who cares about environmental issues.	3.4	1.0
Sustainability is one thing I consider when choosing materials, but it's not the only thing.	3.7	0.9
I prefer materials that look professional, even if they aren't the most sustainable option.	3.2	1.1

Interestingly, the mean rating for "I feel confident I understand the differences between sugarcane paper and white paper" was 3.5, and the mean rating for "I need more information before deciding whether sugarcane paper is a good choice" was 3.4. The similarity in the mean ratings indicates that while participants felt they had some understanding, they acknowledged a need for more information. The high ratings of 3.5 and 3.4 may suggest that a perceived lack of information could be associated with a preference for white paper, along with the other aforementioned cognitive biases. Since participants reported a general lack of information, providing additional information may increase adolescents' openness to sugarcane paper.

Overall, the study's findings suggest that to decrease aversion to non-white, tree-free paper among adolescents, promoting their environmental benefits should be done in conjunction with addressing the cognitive biases that contribute to the preference for white, tree-based paper over non-white, sustainable alternatives. For instance, making the eco-friendly alternatives appear more familiar, popular, or aesthetically pleasing may serve as a useful strategy for the adolescent age group, who will be responsible for leading climate action in the future.

Limitations:

The research design—a cross-sectional study—is a limitation, as although it allows identification of correlations between certain cognitive biases and a preference for white paper, it does not allow the establishment of a causal relationship between the two. As data was collected at a single point in time, it is unclear whether a cognitive bias causes white paper preference, if the preference for white paper strengthens the cognitive bias, or if there exists a bidirectional relationship between the two variables.

The use of opportunity sampling limits the generalizability of the study's findings. As participants were drawn from a single school—a necessary choice due to their prior exposure to sugarcane paper—this sample might not be representative of other student populations or the broader public. Also, demographic information of participants was not collected, which introduces the possibility that the results do not equally represent the perspectives of all year groups. Additionally, because participants were recruited from a school already exposed to sugarcane paper, the findings may not fully reflect responses from populations with no prior exposure to tree-free paper alternatives. Although participants had prior exposure to sugarcane paper through school implementation, white paper remained the more familiar and socially normative material. Therefore, the high Familiarity Bias ratings may reflect relative familiarity rather than complete unfamiliarity with sugarcane paper.

Furthermore, the construct validity of the survey items is a notable limitation. Since no established scales exist to measure cognitive biases in relation to white paper preference, the survey's statements were made by the researcher. Although they were refined multiple times with expert feedback, they have not been externally validated. Hence, it is uncertain whether the statements accurately and distinctly measure all of the

intended cognitive biases. For example, although related, the *Bandwagon Effect* emphasizes adopting beliefs because many others hold them, *Social Proof Bias* emphasizes using others' behaviour as evidence for correctness, and *Appeal to Popularity* concerns judging something as preferable because it is widely liked. This conceptual overlap may have reduced construct validity and construct distinctiveness across some survey items. As a result, some statistically significant differences between cognitive biases may have been less clearly distinguishable conceptually than they appeared numerically. Also, it must be noted that paper preference could also be influenced by variables not measured in this study, such as the affinity for high contrast between black and white or the paper's texture.

The reliance on self-report data collected via a Likert scale survey is also a limitation, as it might have introduced the possibility of social desirability bias, meaning participants may have responded in a manner they felt was socially acceptable and 'environmentally conscious' rather than providing their true opinions. This may have inflated agreement with sustainability-supportive statements or reduced willingness to endorse statements perceived as socially undesirable. Although social desirability bias was inevitable due to the reliance on self-report data, this study minimized this demand characteristic through the use of deception by omission—participants were not informed that the study was investigating cognitive biases—and anonymity.

■ Conclusion

Nevertheless, this study provides a novel psychological framework for understanding the societal preference for white, tree-based paper, a phenomenon highly underexplored due to its pervasiveness. While a purely environmental perspective may assume that awareness automatically leads to action, the psychological approach to sustainability explored by this study recognizes the potential unconscious role of cognitive biases in hindering climate action. Through a survey and an ANOVA, it was found that *Familiarity Bias*, *Appeal to Popularity*, and *Aesthetic Bias* were the cognitive biases most strongly associated with white paper preference. Hence, this research illuminates the importance of addressing the psychological barriers to the adoption of sustainable materials, particularly for current adolescents, who will be the next leaders of climate action. Future research could strengthen these findings by comparing participants with prior exposure to sugarcane paper to participants without such exposure. Another area for future investigation is whether the psychological factors behind white paper preference are replicated in adults, who have more control over the institutional procurement of paper.

■ Acknowledgments

I would like to express my gratitude to Mr. Maddison, the Head of Psychology at my school, for his guidance and support throughout the research process, particularly in ensuring the ethical conduct of this school-wide study.

■ References

- Graham, E. *Office Waste Facts and Trends*. Waste Direct, 2025. <https://wastedirect.co.uk/guides/office-waste-facts-and-trends/> (accessed 2025-05-22).
- Rajas, N.; Kudal, O.; Magar, G.; Marewar, A.; Bilal, M.; Pardeshi, S. Handmade paper formation from onion peels. *International Journal of Membrane Science and Technology* **2023**, *10*(2), 2059–2062. DOI: 10.15379/ijmst.v10i2.2761
- Nunes, L. J. R. The rising threat of atmospheric CO₂: A review on the causes, impacts, and mitigation strategies. *Environments* **2023**, *10*(4), 66. DOI: 10.3390/environments10040066
- Sharma, N.; Bhardwaj, N. K.; Singh, R. B. P. Environmental issues of pulp bleaching and prospects of peracetic acid pulp bleaching: A review. *Journal of Cleaner Production* **2020**, *256*, 120338. DOI: 10.1016/j.jclepro.2020.120338
- Mahmud, Md. A.; Anannya, F. R. Sugarcane bagasse – a source of cellulosic fiber for diverse applications. *Heliyon* **2021**, *7*(8), e07771. DOI: 10.1016/j.heliyon.2021.e07771
- Indie Printing. *Eco-friendly sugarcane paper printing*. Indieprinting.com, 2025. <https://www.indieprinting.com/sugarcane>.
- Han, S. *Welcome to SPRC*. StopPaperRacism, **2025**. <https://www.stoppaperracism.org/>.
- Rachlinks, J. J. *The psychology of global climate change*. Scholarship@Cornell Law: A Digital Repository, **2025**. <https://scholarship.law.cornell.edu/facpub/792>.
- Korteling, Johan. E., Paradies, G. L.; Sassen-van Meer, J. P. Cognitive bias and how to improve sustainable decision making. *Frontiers in Psychology* **2023**, *14*(1129835). DOI: 10.3389/fpsyg.2023.1129835
- Kahneman, D. *Thinking, fast and slow*. Farrar, Straus and Giroux, 2011.
- Dhami, S.; Hajimoladarvish, N.; Mamidi, P. Climate change risk, and human behavior: Theory and evidence. *SSRN Electronic Journal* **2023**. DOI: 10.2139/ssrn.4592965
- Direr, A. Bringing present bias back to the present. *SSRN Electronic Journal* **2019**. DOI: 10.2139/ssrn.3420916
- Liu, Y. The review of loss aversion. *Advances in Education Humanities and Social Science Research* **2023**, *7*(1), 428. DOI: 10.56028/aehtsr.7.1.428.2023
- Cheng, Z. The impact of status quo bias on people's choices. *Advances in Economics, Management and Political Sciences* **2023**, *10*(1), 111–115. DOI: 10.54254/2754-1169/10/20230442
- Liebe, U.; Gewinner, J.; Diekmann, A. What is missing in research on non-monetary incentives in the household energy sector? *Energy Policy* **2018**, *123*, 180–183. DOI: 10.1016/j.enpol.2018.08.036
- Liebe, U.; Gewinner, J.; Diekmann, A. Large and persistent effects of green energy defaults in the household and business sectors. *Nature Human Behavior* **2021**, *5*(5), 576–585. DOI: 10.1038/s41562-021-01070-3
- Sunstein, C. R.; Reisch, L. A. Climate-friendly default rules. *Sustainable Consumption and Production* **2021**, *1*, 141–164. DOI: 10.1007/978-3-030-56371-4_8
- Han, S. *Paper Preferences and Perceptions* [Survey]. <https://forms.gle/gcMLbVdV2Q6fdUFA>.
- Bullet, P. B. *How to choose a sample size (for the statistically challenged)*. <https://tools4dev.org/resources/how-to-choose-a-sample-size/>.
- Pinho, A. da S.; Molleman, L.; Braams, B. R.; van den Bos, W. Majority and popularity effects on norm formation in adolescence. *Scientific Reports* **2021**, *11*(1), 12884. DOI: 10.1038/s41598-021-92482-8
- Somerville, L. H.; Jones, R. M.; Ruberry, E. J.; Dyke, J. P.; Glover, G.; Casey, B. J. The medial prefrontal cortex and the emergency

- of self-conscious emotion in adolescence. *Psychological Science* **2013**, 24(8), 1554-1562. DOI: 10.1177/0956797613475633
22. Van Petegam, S.; Soenens, B.; Vansteenkiste, M.; Beyers, W. Rebels with a cause? Adolescent defiance from the perspective of reactance theory and self-determination theory. *Child Development* **2015**, 86(3), 903-918. DOI: 10.1111/cdev.12355
23. Fishbach, A.; Eskreis-Winkler, L. Hidden failures. *Academy of Management Proceedings* **2018**, 2018(1), 14861. DOI: 10.5465/ambpp.2018.292

■ Author

Seoha Han is a student researcher who combines psychological research with efforts to change unsustainable systems. She looks forward to continuing her interdisciplinary research at university and beyond.

CRISPR-Cas9 for Treating Genetic Blood Disorders: Promises and Limits

Shriyans Katukuri

Interlake High School 16245 NE 24th St, Bellevue, Washington, 98008, USA; shriyans.katukuri@outlook.com

ABSTRACT: Gene editing is a modern tool that shows promise in the treatment of genetic disorders. This review discusses major improvements in CRISPR regarding its mechanisms, delivery, safety, and clinical use. CRISPR alters DNA sequences to alter a gene's function. This is usually done by editing patient stem cells or immune cells through delivery methods. The clearest human efficacy signal comes from hemoglobinopathies, where editing increases production of protective fetal hemoglobin or corrects β -globin biology, leading to proven clinical benefit. In oncology, CRISPR's role is to enable more resilient cell therapies—by standardizing insertion sites and removing inhibitory checkpoints—rather than directly rewriting heterogeneous tumor genomes. Targeted delivery, preservation of stemness, genomic integrity, immune recognition of editing components, manufacturing complexity, and equitable access are examples of questions that surround CRISPR. New studies address these concerns with their inclusion of transient editors, double-strand-break-sparing technologies, orthogonal off-target assessments, and rigorous lot-release criteria. However, while these new studies are furthering CRISPR studies in the right direction, longer follow-up is still needed. Overall, the field has shifted from theory to clinical reality. CRISPR's next steps should revolve around targeting possible scientific improvements and prioritizing ethical use and access to the tool.

KEYWORDS: Cell Immunology, Cell and Tissue Engineering, Genetics, Genetic and Molecular Biology of Disease, Disease Treatment and Therapies.

■ Introduction

DNA is the building block that determines the composition of every living being, but small disruptions in this code can often lead to major health problems. Many genetic disorders—such as cystic fibrosis, sickle cell disease (SCD), and β -thalassemia—are caused by mutations in DNA that affect a gene's function.¹ Traditional treatments have only been able to target the symptoms of these disorders through methods such as bone marrow transplants, transfusions, and hydroxyurea rather than directly targeting their underlying genetic causes.² Despite countless years of research, there have been very few slightly curative solutions, leaving an issue in genetic medicine. There is a lack of an accessible, deliberate, and durable solution to correct pathogenic mutations at their root source.

In modern-day science, advances in molecular biology have begun to target this issue. Initial genome editing tools, such as zinc-finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs), have demonstrated that human DNA can be modified, but these tools were too expensive, complex, and difficult to utilize to target each new mutation.³ The discovery of CRISPR–Cas9 has revolutionized the field by providing a simpler, more programmable tool for site-specific DNA editing.^{2,3} With the technology of programming RNA guides with enzymes, CRISPR has been able to target the root cause of several genetic disorders. This technology has given the scientific world the ability to repair pathogenic mutations in stem cells, allowing for a multitude of clinical possibilities to open up. Yet, as CRISPR–Cas9 advances and moves from preclinical research into human application, it becomes increasingly important to understand the technology's promises

and its limitations. Identifying CRISPR–Cas9's abilities, such as allowing durable correction of monogenic mutations in accessible cell populations like hematopoietic stem cells, allows scientists to utilize the technology safely. Additionally, understanding the technology's limitations, such as its inability to target a plethora of genes without great cost, allows scientists to navigate the liabilities needed to translate clinical successes into safe, scalable, and ethical therapies.

As it stands, CRISPR is good at targeting single genes to attack a malicious tumor, while it falls short in terms of being able to edit complex genetic mutations.^{4,5} This review argues that CRISPR–Cas9 has already achieved clinically durable benefit in monogenic hemoglobinopathies through *ex vivo* editing of hematopoietic stem and progenitor cells (HSPCs), but maintaining similar successes in polygenic blood cancers and improving the accessibility of CRISPR requires improvements in CRISPR itself, such as biological improvements and overall improved regulation of the tool.

■ Methodology

This is a narrative, problem-oriented literature review covering 2010–2025, the period in which the first reports of CRISPR–Cas9 editing in mammalian cells were reported, to the first regulatory approvals and late-phase results. PubMed, Google Scholar, and ClinicalTrials.gov were the primary sources from which journals and studies were obtained. A significant number of regulatory reports from the FDA/EMA were cross-referenced with sources and trials to check for approval status and safety monitoring.

Key terms such as "CRISPR–Cas9 hematopoietic stem cells," "BCL11A enhancer editing," "fetal hemoglobin re-activation," "CTX001/exa-cel results," "base editing HBG promoters," "CRISPR off-target," "high-fidelity Cas9," "electroporation," and "CRISPR-engineered immune cells," etc., were used in the process of searching for reliable sources. Sources that contained well-characterized editing strategies and readouts (on-target rate, off-target profiling, functional rescue), clinically meaningful results in human trials (e.g., transfusion independence, VOC-free survival), or insight with applicability to blood disorders were prioritized in the source selection process. Non-academic sources were only included to verify the clinical usage of gene editing tools and to back up common claims, and peer-reviewed studies and governmental recognition verified all data from such sources. It is important to note that all sources were gathered online without any physical tools or materials.

Screening was done in three phases. Firstly, papers outside hematology or primary data were excluded. Second, trials that had rigorous specificity testing (e.g., GUIDE-seq), new enzyme innovations (e.g., eSpCas9, SpCas9-HF1), and good delivery/manufacturing methods were prioritized. Finally, the results were sorted by the topics that they discussed. Basic CRISPR mechanisms where editing is most useful (monogenic vs. complex disease); delivery methodologies (*ex vivo* electroporation, AAV6 donors, lipid nanoparticles), safety (off-targets, large on-target changes, immunity), and clinical translation (exa-cel results; building base-editing programs) were all topics of major interest.

Brief History and Mechanism of Gene Editing:

Gene editing has a long history that did not begin with CRISPR. Early tools, such as zinc finger nucleases (ZFNs) and transcription activators like effector nucleases (TALENs), allowed scientists to cut DNA at chosen sites. However, these technologies were hard to design and manipulate to work for each new target.^{2,5,6} CRISPR–Cas9 changed the field by making targeting DNA simple. You swap a short guide RNA to point CRISPR–Cas9 to almost any DNA sequence, and then CRISPR–Cas9 makes a double-strand break at that site.⁵ The cell then fixes that break by one of two main pathways. The first being non-homologous end joining (NHEJ).⁴ NHEJ is fast, but it often isn't accurate.⁵ It was common for NHEJ to delete or insert new letters and completely change a gene and cause mutations.^{4,5}

Homology-directed repair (HDR) is slower but allows scientists to cause precise changes when a DNA repair template is provided.^{4–6} The first landmark papers in 2012–2013 showed that CRISPR–Cas9 could work in bacteria and mammalian cells while allowing for multiplex editing, which meant that several genes could be changed at once.^{2,3,7} Researchers then began mapping off-target cutting and building "high-fidelity" Cas9 variants to improve specificity in gene targeting and editing.⁴ At the same time, gene editing technologies delivery went from viruses to electroporation and lipid nanoparticles, depending on the tissue.^{7,8}

Together, these advances have turned gene editing from a niche lab technique into a mainstream platform with real clinical paths. Due to the biology of blood disorders, CRISPR is able to target their root causes with more ease compared to other types of disorders.⁹ Blood stem cells can be collected, edited outside the body, where quality can be checked, and then reinfused after conditioning.⁹ The practicality of this, plus strong genetic targets, explains why hematology became the first clinical proving ground for CRISPR.

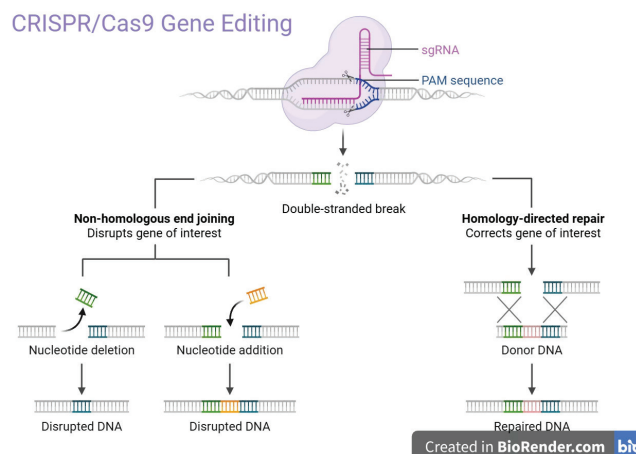


Figure 1: CRISPR–Cas9 Process. This figure shows the process of CRISPR–Cas9 gene editing with either non-homologous end joining or homology-directed repair. Created using BioRender.

Introduction to CRISPR–Cas9: Mechanism and Significance:

CRISPR–Cas9 is a two-part system. There is the guide RNA that finds a matching DNA sequence, and the Cas9 enzyme that cuts there, as shown in Figure 1.⁶ This can be thought of as the "address" of the target and Cas9 as the "scissors" that edit the target. Because the guide is easy to design, CRISPR can be retargeted quickly across the genome.⁴ After Cas9 makes a cut, cells repair the change either by NHEJ (which often disables the gene) or HDR (which allows for precise fixes), as illustrated in Figure 1.⁷ Base editors chemically change a single DNA letter without cutting both strands, which can lower the risk of big changes to the DNA sequence.¹⁰ Prime editors use a reverse transcriptase to "write" new sequences with fewer breaks.^{6,10}

In the blood, editing is typically done *ex vivo* by electroporating CRISPR components into hematopoietic stem and progenitor cells (HSPCs) and then transplanting them back.¹¹ For liver diseases, lipid nanoparticles deliver CRISPR *in vivo*.¹² This approach was proven to work well in transthyretin amyloidosis.¹² Any accidental cuts are tracked with unbiased methods such as GUIDE-seq to eliminate unintentional edits caused by the editing.^{4,7} The next-generation Cas9 variants significantly reduce spurious activity that was common in previous generations of the tool.²

While CRISPR seems to work in almost all use cases, scientists also must pay attention to immune responses against bacterial Cas9 proteins and to large off-target deletions that sometimes arise from double-strand break repair.^{10,13,14} CRISPR matters because it mixes precision with practicality,

allowing for better treatment of genetic disorders. For blood disorders that have clear genetic drivers, it lets us either correct the mutation directly or change a regulatory switch that compensates for it, with results already strong enough to earn the early regulatory approvals that are already saving lives.^{9,15}

Monogenic vs. polygenic blood disorders — where CRISPR fits best:

CRISPR works best when the target it's dealing with is simple and is vital to the creation of the disorder.³ Single-gene (monogenic) diseases like sickle cell disease (SCD) and transfusion-dependent β -thalassemia (TDT) fit that category.⁹ Well-understood changes in the β -globin locus drive both of them.⁹ Editing a regulatory element to raise fetal hemoglobin (HbF) can bypass the defective adult hemoglobin and reduce symptoms in the disorder. Clinical trials with this strategy, as seen by the Exagamglogene Autotemcel (exa-cel/Casgevy) trials, show durable HbF induction, transfusion independence in TDT, and elimination of Vaso-occlusive crises that occur in most SCD patients.^{9,15}

Polygenic or genomically complex diseases like acute lymphoblastic leukemia (ALL) are tougher to combat with CRISPR because ALL involves a plethora of mutations and clones that evolve.¹⁶ That makes a single "fix" unrealistic. Instead, CRISPR's value in curing the disorder is often indirect. CRISPR-Cas9 engineers immune cells (like CAR-T cells) to work better; or by modeling resistance pathways or targets; and by combining edits (such as knocking out PD-1 or TCR chains) to kill a greater number of tumors with great effectiveness while limiting exhaustion.¹⁶

In short, monogenic blood disorders are the easiest to attack with CRISPR because CRISPR only has to edit one gene in order to be effective in removing the disorder. In polygenic disorders, specifically leukemias, CRISPR is more of a tool to assist the current cell therapies and to study disease genetics rather than being the final cure-all solution.

Snapshots of key blood disorders with single-gene drivers:

Many hematologic diseases are monogenic, making them natural candidates for CRISPR intervention, while others are syndromic or in predisposition states where editing may be feasible but ethically and technically complex.

An example of such monogenic diseases would be Hemoglobinopathies (*HBB* locus). Sickle cell disease (SCD) and transfusion-dependent β -thalassemia (TDT) are caused by pathogenic variants in *HBB*. Editing either one reactivates fetal hemoglobin by disrupting the erythroid *BCL11A* enhancer or by altering *HBB* promoters. *Ex vivo* CRISPR tools have delivered durable HbF induction with the removal of Vaso-occlusive crises in SCD and transfusion independence in most TDT patients, leading to regulatory approval.^{3,15,17}

Further diseases, such as Hemophilia A/B (*F8/F9*), are also monogenic diseases. Hemophilia A results from *F8* variants and Hemophilia B from *F9* variants, both X-linked genes and characterized by factor VIII or IX deficiency and bleeding diathesis.¹⁸ CRISPR strategies aim at editing the liver (e.g., AAV/LNP-delivered editors) to restore factor production and

correct hotspot mutations. With this disorder, base/prime editing may reduce double-strand break risks.¹⁸

Inherited neutropenia is a monogenic blood disorder where pathogenic variants in *ELANE* (neutrophil elastase) can lead to cyclic or severe congenital neutropenia, with a high risk of infection and a measurable risk of MDS/AML in the congenital form.¹⁹ Gene correction in autologous HSCs is the logical path to treatment, but it contains its dangers due to the chance that malignant mutations can occur independent of therapy.

Another monogenic blood disorder would be bone marrow failure/ribosomopathies. Diamond-Blackfan anemia is most often caused by haploinsufficiency, where a patient only has one functional copy of a gene, for ribosomal protein genes (e.g., *RPS19*), producing red cell aplasia with congenital anomalies.²⁰ The genetic heterogeneity (dozens of genes) complicates a single target edit, though base editing of specific founder alleles is being studied preclinically.⁵

Preclinical gene therapy groundwork in SCD and TDT:

Before the first human CRISPR trials, laboratory research demonstrated that genetic editing could correct the sickle cell mutation in human CD34+ cells and recreate healthy red blood cells.²¹ Early research used HDR templates to fix the Glu6Val substitutions in *HBB*, leading to a restoration in adult hemoglobin levels in edited cells that could then be grafted into mice.^{21,22} Concurrently, researchers discovered another method of correcting the sickle cell mutation. Instead of editing *HBB* directly (which is hard to scale with HDR), if a regulatory switch is disrupted, the fetal hemoglobin levels will rise. Through saturation mutagenesis, the small enhancer of the *BCL11A* gene in erythroid cells was mapped, and with that, scientists learned that editing this enhancer sharply increases HbF, which does not cause red blood cells to be sickled and can compensate for *HBB* defects.¹⁷

This strategy is the more commonly used one because it relies on NHEJ (efficient) and because HbF reactivation has a long clinical track record of success in hydroxyurea-treated patients. These preclinical successes led to the *ex vivo* design being used clinically.²³ The process is as such. First, HSPCs are mobilized, then CRISPR-Cas9 RNPs that target the *BCL11A* erythroid enhancer (or nearby HbF promoters) are electroporated, the product is verified and double checked to account for concerning off-targets, and finally, the product is reintroduced into the patient's body.¹⁷ Animal models further confirmed long-term engraftment and normal HbF expression from edited stem cells. By 2020–2024, data from multicenter trials showed transfusion independence in TDT and near complete suppression of Vaso-occlusive crises in SCD, paving the way for widespread use of this method to combat SCD.^{3,15} These studies also helped create the safety guidelines in these types of treatments.²³ By creating a process where off-targets are monitored with unbiased assays, large deletions are quantified and measured, and bank backup cells are stored in case re-collection is needed.

Why *BCL11A* enhancer targeting works (and other promising targets):

BCL11A represses γ -globin in adult-type red cells. In people with certain natural mutations in the γ -globin genes, HbF levels stay high for the person's lifetime, and their cells are protected against SCD and TDT. CRISPR leverages this biology by disrupting a short, erythroid-specific enhancer within intron 2 of *BCL11A*. Because that enhancer is active primarily in red blood cell precursors, editing it raises HbF levels without removing *BCL11A* in the immune or nervous systems, where the gene has other important roles.¹⁷

Clinically, gene enhancer edits lift HbF levels to a level that dilutes or replaces sickling hemoglobin, which leads to improving anemia and preventing crises in the blood. Alternative targets follow the same logic. By editing the *HBG1/2* promoters to weaken *BCL11A* binding (base editing with BEAM-101), or by tweaking other transcriptional regulators in the fetal hemoglobin to adult hemoglobin switch.²⁴ Outside hemoglobinopathies, targets depend on disease biology. In immune cells, edits to *TRAC* (to standardize CAR expression), edits to *PDCD1* (PD-1, to sustain activity), or edits to checkpoint pathways can boost antitumor effects in cells.^{25,26}

In leukemias, some groups test the editing of surface antigens or resistance nodes, but the difficulties are greater when changing malignant genomes.¹⁶ Overall, *BCL11A* enhancer targeting works due to its nature, which is mechanistically grounded, efficient with NHEJ, and tissue-restricted, which explains its success compared with more invasive genome rewrites. Base editing and prime editing may improve treatments further by avoiding double-strand breaks and reducing risks like large DNA sequence deletions.

Emerging trials: *exa-cel* (CTX001/Casgevy) and next-gen editors:

The most advanced gene editing program as of right now is Exagamglogene Autotemcel (*exa-cel*; formerly CTX001, now Casgevy), which edits the *BCL11A* erythroid enhancer in autologous HSPCs.^{9,15,27} In pivotal studies, most SCD patients became free of Vaso-occlusive crises, and most TDT patients achieved durable transfusion independence by using this therapy.¹⁵ The U.S. and U.K. authorized the therapy in late 2023, allowing for it to become mainstream.^{8,10,27} Now, base editors aim to raise HbF levels while avoiding double-strand breaks. BEAM-101 base edits *HBG* promoters to reduce *BCL11A* bindings with early reports showing a strong HbF induction (above 60%), rapid engraftment, and no post-engraftment crises in small trial cohorts, though data is still being collected.²⁴

Editas's *Revi-cel* (formerly EDIT-301) uses an AsCas12a nuclease to cut *HBG* promoters. Early updates in the trial also suggest high HbF levels and crisis freedom with limited follow-ups for patients. For cancers, CRISPR enables "armored" T cells and precise CAR attacks at defined loci like *TRAC*, which produce more uniform, less exhausted cells in preclinical models, and the first human polygene editing has shown proof that it may be successful.^{25,26} Trials conducted, such as this one focusing on NTLA-2001, which uses lipid nanoparticles delivering Cas9 mRNA and guide RNA to the liver,

reduce transthyretin protein levels in patients with ATTR amyloidosis and validated systemic editing in humans.^{5,28}

Together, these studies outline a future with tailored editors (Cas9, Cas12, base, prime), smarter cell therapies, and broader applications, given that long-term safety in gene editing remains possible.

Delivery methods – viral vectors, nanoparticles, and electroporation:

The method of delivery matters just as much as the genetic edit itself. For blood disorders, the standard approach of delivery is to edit cells outside the body (*ex vivo*). Doctors collect a patient's blood stem cells (CD34⁺ HSPCs), then use electroporation to briefly open the cell membrane and push in a CRISPR-Cas9 guide RNA ribonucleoprotein (RNP).²⁹ Because the nuclease is present only for minutes to hours, this reduces off-target risk and avoids long-term enzyme expression.^{3,30}

If a precise insert is needed, such as placing a CAR at the *TRAC* site, a short DNA "template" will be supplied during the repair, with AAV6 often being used to deliver that template for homology-directed repair (HDR).²⁵ Many groups avoid viral delivery of the nuclease itself when used, as AAV6 is typically just a donor for HDR, not the editor.³ *In vivo* delivery is different as the solution is directly inserted into the target region. Lipid nanoparticles (LNPs) that carry Cas9 mRNA and guide RNA to the liver enabled the first systemic CRISPR editing in people (NTLA-2001), demonstrating that infusion-based editing is feasible.²⁸ As it stands, bone-marrow targeted LNPs are under active study, but most hematology programs still use *ex vivo* editing because it's auditable and allows strict lot release testing before cells are returned to the patient.^{3,12} Quality control typically checks editing efficiency, off-targets, vector copies (if any), and cell function (stemness for HSPCs, killing capacity for T cells). As base and prime editors mature and the methods of delivery improve, more *in vivo* hematology applications may appear, but the current clinical standard remains clean, virus-free *ex vivo* manufacturing.^{12,31}

Safety concerns – off-targets, immune responses, and large edits:

Most major safety concerns regarding CRISPR-Cas9 can be grouped into three main groups. First, off-target edits. RNA guides can sometimes send Cas9 to similar DNA sites, leading to further damage from CRISPR.⁷ Researchers lower this risk with improvements in guide design, high-fidelity Cas9 variants, and unbiased mapping tools like GUIDE-seq before products reach patients.³²

Second, immune responses. Because Cas9 comes from bacteria, some people already have T cells or antibodies that recognize it.³⁰ This is one of the main factors as to why *ex vivo* products use transient RNPs. Early trials have shown that edited cells can persist without obvious immune rejection, disproving this concern.³³

Third, large on-target changes. Double-strand breaks can occasionally cause big deletions or rearrangements at the intended cut site that basic PCR might miss. Long-read sequencing and

other assays have detected such events in HSPCs and T cells at low to moderate rates depending on the locus.¹⁰

Doctors work to minimize double-strand breaks (short Cas9 exposure, high-fidelity enzymes), use DSB-free editors when a single-letter change will do (base/prime), and add release assays to screen out risky alleles to alleviate these concerns.^{10,20,33} For cell therapies, the usual CAR-T toxicities (cytokine release syndrome and neurotoxicity) still apply and are monitored as standard of care.²⁶

Table 1: Collection of clinical trials from clinicaltrials.org, evaluating gene editing therapies in select blood disorders.

Product / Name	Condition(s)	Editing modality	Genetic target	Key Findings
Exagamglogene Autotemcel (exa-cel, CTX001) - CLIMB THAL-111	Transfusion-dependent β -thalassemia(TDT)	CRISPR-Cas9	<i>BCL11A</i> erythroid enhancer	Most patients achieved durable transfusion independence; Hb typically >11 g/dL; safety consistent with busulfan conditioning. ³⁴
Exagamglogene Autotemcel (exa-cel, CTX001) - CLIMB SCD-121	Sickle cell disease (SCD)	CRISPR-Cas9	<i>BCL11A</i> erythroid enhancer	97% of evaluable pts free from severe VOCs for ≥ 12 months; HbF ~40% with pancellular distribution; safety consistent with conditioning. ³⁵
CLIMB-131 long-term follow-up (LTFU)	SCD and β -thalassemia long-term follow-up of CTX001 recipients	CRISPR-Cas9	<i>BCL11A</i> erythroid enhancer (parent studies)	Edited allele levels stable in blood & marrow; sustained clinical benefit reported in follow-up cohorts. ³⁶
EDIT-301 (Reni-cel) - RUBY	Severe sickle cell disease	CRISPR-Cas12a (Cpf1)	<i>HBB</i> /1/2 promoters (y-globin upregulation)	Early data show robust HbF (~40-50%), total Hb ≥ 11 g/dL, and elimination of VOCs in most evaluable pts to date. ³⁷
EDIT-301 - EditHAL	Transfusion-dependent β -thalassemia	CRISPR-Cas12a (Cpf1)	<i>HBB</i> /1/2 promoters	Early participants show increased HbF and reductions in transfusion needs; some achieve transfusion independence. ³⁸
BEAM-101 - BEACON	Severe sickle cell disease with recurrent VOCs	CRISPR adenine base editing	<i>HBB</i> /1/2 promoter	Initial cohorts show increased HbF (>40% in many) and reductions in VOCs; safety is acceptable so far. ³⁹
BIVV003 - PRECIZN-1	Severe sickle cell disease	Zinc Finger Nuclease	<i>BCL11A</i> erythroid enhancer	Early pts showed HbF elevation (e.g., ~45% in one improved-manufacture pt) and VOC reduction; program deprioritized pending partner. ⁴⁰
Long-term safety/efficacy follow-up for BIVV003 / ST-400	LTFU in SCD (BIVV003) & β -thalassemia (ST-400) recipients	ZFN	<i>BCL11A</i> enhancer (parent)	LTFU safety/efficacy monitoring of prior ZFN-treated participants. ⁴¹
ST-400 - THALES	Transfusion-dependent β -thalassemia	Zinc Finger Nuclease	<i>BCL11A</i> erythroid enhancer	Prompt engraftment; HbF induction observed; mixed clinical benefit; limited dataset. ⁴²
SB - FIX	Hemophilia B	Zinc Finger Nuclease	Albumin locus knock-in of <i>F9</i>	First-in-human <i>in vivo</i> ZFN editing; no sustained FIX expression; program discontinued. ⁴³
OTQ923	Sickle cell disease	CRISPR-Cas9	<i>BCL11A</i> (HbF induction)	Proof-of-concept <i>BCL11A</i> -targeting program; study ended; no posted outcomes on registry. ⁴⁴
Exa - cel pediatric - CLIMB - 151	Pediatric severe SCD (ages 2-11)	CRISPR-Cas9	<i>BCL11A</i> erythroid enhancer	Pediatric expansion of exa-cel; results pending. ⁴⁵
Exa - cel in HbSC	Severe SCD, B α /B β genotype (HbSC)	CRISPR-Cas9	<i>BCL11A</i> erythroid enhancer	Early-stage CRISPR program; no human efficacy data yet. ⁴⁶
CS - 101 (CorrectSequence)	Severe sickle cell disease	Base editing	HBB promoter (<i>BCL11A</i> binding site)	Early-stage CRISPR program; no human efficacy data yet. ⁴⁷

■ Discussion

This collection of trials revolving around monogenic blood disorders shows that, as discussed before, the CRISPR-Cas9's efficacy is strongly backed by the evidence from the results discovered by these trials, as summarized in Table 1. Looking at the table, the *ex vivo* editing of autologous CD34⁺ hematopoietic stem and progenitor cells (HSPCs) has consistently delivered durable clinical benefit in both β -thalassemia and sickle cell disease (SCD).^{9,15,24} In transfusion-dependent β -thalassemia, Exagamglogene Autotemcel (exa-cel; Casgevy) treatment led to transfusion independence in around 91% of patients, with mean hemoglobin concentrations around 13 g/dL and pan-cellular HbF distribution.¹⁵ Similarly, in sickle cell disease, roughly 97% of participants were free from severe

Vaso-occlusive crises (VOCs) for more than 12 consecutive months, and 100% avoided VOC-related hospitalizations over the same time period, with sustained increases in HbF and improvements in hemolysis markers.^{3,21,22} Table 1 highlights a consistent association between CRISPR-based editing and increases in fetal hemoglobin. These findings, as well as regulatory approval for public use by the U.S Food and Drug Administration in 2023, give reason to believe that *ex vivo* CRISPR-Cas9 editing of CD34⁺ HSCs can be used commercially to produce persistent benefits in humans.²⁷

However, translating these same clinical successes to blood cancers requires that the role of CRISPR be more indirect. Rather than directly repairing malignant genomes—which are often unstable and polyclonal—CRISPR is used to enable the precise engineering of immune effector cells.^{14,33} For example, targeted integration of chimeric antigen receptor (CAR) constructs at the *TRAC* locus reduces tonic signaling, promotes uniform CAR expression, and enhances antitumor persistence in acute lymphoblastic leukemia.^{16,25} Multiplex editing has allowed for the simultaneous removal of inhibitory receptors such as PD-1 and TCR chains, producing armored T cells with improved resilience and reduced cell fatigue.²⁶ In pivotal studies on anti-CD19 CAR-T therapies—such as tisagenlecleucel and Brexucabtagene autotemcel—continue to set clinical benchmarks, with complete response (CRi) rates around 70%, drastically improving survival outcomes in patients.⁴⁸ CRISPR-enabled T-cell engineering contains the potential to eventually standardize manufacturing, minimizing variability between batches of solutions, and overall lead to “off the shelf” allogeneic cell therapies that can be used as regular medicine that can be more widely accessible.

While the promise of allogeneic CRISPR solutions is one of success and feasibility, the transition from autologous to allogeneic CRISPR platforms will require careful evaluations of individual immune compatibility and graft versus host risks.³⁰⁻³² Nuanced approaches to these issues, such as multiplex knockouts of *HLA* genes or the additions of edits to improve the stealth of edited cells to evade a host's immune system, are currently being explored. At the moment, ensuring the genomic stability of heavily edited cells is a priority in research regarding CRISPR. Long-term sequencing and orthogonal mapping methods have detected multiple low-frequency large deletions in DNA sequences and rearrangements at target loci, which might carry unpredictable long-term consequences.¹⁰ Therefore, there must be a balance between the ambition of multi-gene editing and the safety guidelines that have been established for clinical trust.

While CRISPR technology may be feasible scientifically, there are significant discussions regarding the ethical aspects of CRISPR usage. While Exa-cel may have received authorization to be used in the market, the therapy's estimated list price of around 2.2 million USD illustrates the unethical divide between members of differing socioeconomic groups, determining access to life-saving treatments.⁴⁹ Even with managed access agreements, treatments are restricted to a small number of specialized centers in high-income countries. For low and middle-income nations where hemoglobinopathies are most

common, these economic barriers are prohibitive. If CRISPR is to be fully realized as what it is meant to be, scientists must scale manufacturing, simplify conditioning regimens, and build global infrastructure for genomic therapies. While the ethical concerns of access exist, there is a much larger debate over whether genome editing is a moral practice. Both the U.S. National Academies (2017) and the WHO (2021) emphasize that somatic genome editing should only be conducted with strict frameworks in the usage of this tool, with no margin for error.^{50,51} In order to practically implement this technology, there should be rigorous informed consent procedures, community engagement, and global data sharing to track the long-term outcomes of genome editing. As the field shifts towards *in vivo* editing approaches and base or prime editors that blur the line between somatic and heritable changes, there must be ethical oversight over this whole process. With CRISPR having the ability to not only edit the genes of one individual but also their bloodline, genome editing should be used with care.

Finally, while CRISPR's success in hemoglobinopathies is evident, the next step for the technology involves extending its reach. This includes refining delivery systems such as bone marrow tropic lipid nanoparticles, exploring gentler conditioning methods, and applying double-strand break-free editors for more predictable outcomes. CRISPR's future now depends not only on molecular innovation but on the whole field of bioengineering, regulatory policy, and social equity.

■ Conclusion

In conclusion, CRISPR-Cas9 has proven its reliability in delivering consistent, durable clinical benefits when treating monogenic blood disorders, specifically sickle cell disease (SCD) and transfusion-dependent β -thalassemia (TDT). Furthermore, when the results from the trials were analyzed, the *ex vivo* editing of cells led to positive results being found within patients. With these results being combined with the fact that some CRISPR-Cas9 treatments, such as exa-cel, have received regulatory authorization in the United States and the United Kingdom, CRISPR-Cas9 establishes itself as one of the leading modern tools in the battle against genetic disorders.

In polygenic blood disorders, CRISPR shows potential in the treatment of these genomically complex malignancies, such as leukemias and lymphomas, by utilizing its ability to adapt to a multitude of scenarios. With researchers using CRISPR-Cas9 to improve immune systems and support other systems within the body, CRISPR-Cas9 can still fight against polygenic disorders through unconventional routes. Site-specific integration of chimeric antigen receptors (CARs) at the *TRAC* locus and the multiplex deletion of inhibitory receptors like TCR chains enhance the ability of CAR-T cell therapies.

Finally, CRISPR-Cas9 has advanced from being a theoretical concept to a critical clinical tool. From its success in monogenic disorders to its adaptability in the treatment of polygenic blood disorders, the evidence shows that genome editing can produce sustainable cures when used with precision, safety, and ethical oversight. As the various aspects of CRISPR, such as its delivery methods and global accessibility,

continue to be improved, CRISPR will redefine how medicine approaches the treatment of blood disorders.

While this is a comprehensive review regarding the subject of CRISPR and blood disorders, there is a plethora of limitations. Firstly, it must be recognized that while there are a multitude of sources within this review, this review cannot include and compare all the sources available regarding this subject. The decision in choosing which sources to reference and use in the review may have led to some information biases. The rapid nature of this field, in which discoveries are constantly being made, leads to this review being unable to account for current discoveries from Aug 25th, 2025, and onwards. Finally, due to many of the trials being analyzed still being continued as those trials mature, the conclusions established in this review may need to be revisited.

■ Acknowledgments

I would like to thank Dr. Prahlad Parajuli for helping me with the scientific aspect of this journal and for helping me with the overall process. I would also like to thank Professor Kay Diaz and Professor Virgel Torremocha for their help in the editing process of this review.

■ References

- Genetic Disorders: *What Are They, Types, Symptoms & Causes*. Cleveland Clinic. <https://my.clevelandclinic.org/health/diseases/21751-genetic-disorders?> (accessed 2025-11-06).
- Porteus, M. H. Towards a New Era in Medicine: Therapeutic Genome Editing. *Genome Biology* **2015**, *16* (1). <https://doi.org/10.1186/s13059-015-0859-y>.
- Lino, C. A.; Harper, J. C.; Carney, J. P.; Timlin, J. A. Delivering CRISPR: A Review of the Challenges and Approaches. *Drug Delivery* **2018**, *25* (1), 1234–1257. <https://doi.org/10.1080/10717544.2018.1474964>.
- Mali, P.; Yang, L.; Esvelt, K. M.; Aach, J.; Guell, M.; DiCarlo, J. E.; Norville, J. E.; Church, G. M. RNA-Guided Human Genome Engineering via Cas9. *Science* **2013**, *339* (6121), 823–826. <https://doi.org/10.1126/science.1232033>.
- Cong, L.; Ran, F. A.; Cox, D.; Lin, S.; Barretto, R.; Habib, N.; Hsu, P. D.; Wu, X.; Jiang, W.; Marraffini, L. A.; Zhang, F. Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* **2013**, *339* (6121), 819–823. <https://doi.org/10.1126/science.1231143>.
- Savić, N.; Schwank, G. Advances in Therapeutic CRISPR/Cas9 Genome Editing. *Translational Research* **2016**, *168* (26470680), 15–21. <https://doi.org/10.1016/j.trsl.2015.09.008>.
- Kleinstiver, B. P.; Pattanayak, V.; Prew, M. S.; Tsai, S. Q.; Nguyen, N. T.; Zheng, Z.; Joung, J. K. High-Fidelity CRISPR–Cas9 Nucleases with No Detectable Genome-Wide Off-Target Effects. *Nature* **2016**, *529* (7587), 490–495. <https://doi.org/10.1038/nature16526>.
- Wang, D.; Zhang, F.; Gao, G. CRISPR-Based Therapeutic Genome Editing: Strategies and *in Vivo* Delivery by AAV Vectors. *Cell* **2020**, *181* (1), 136–150. <https://doi.org/10.1016/j.cell.2020.03.023>.
- Handgretinger, R.; Mezger, M. An Evaluation of Exagamglogene Autotemcel for the Treatment of Sickle Cell Disease and Transfusion-Dependent Beta-Thalassaemia. *Expert Opinion on Biological Therapy* **2024**, *24* (9), 883–888. <https://doi.org/10.1080/14712598.2024.2399134>.
- Kosicki, M.; Tomberg, K.; Bradley, A. Repair of Double-Strand Breaks Induced by CRISPR–Cas9 Leads to Large Deletions and

- Complex Rearrangements. *Nature Biotechnology* **2018**, *36* (765–771). <https://doi.org/10.1038/nbt.4192>.
11. Testa, U.; Leone, G.; Cappellini, M. D. Therapeutic Gene Editing for Hemoglobinopathies. *Mediterranean journal of hematology and infectious diseases* **2024**, *16* (1), e2024068. <https://doi.org/10.4084/MJHID.2024.068>.
 12. Gillmore, J. D.; Gane, E.; Taubel, J.; Kao, J.; Fontana, M.; Maitland, M. L.; Seitzer, J.; O'Connell, D.; Walsh, K. R.; Wood, K.; Phillips, J.; Xu, Y.; Amaral, A.; Boyd, A. P.; Cehelsky, J. E.; McKee, M. D.; Schiermeier, A.; Harari, O.; Murphy, A.; Kyrtatous, C. A. CRISPR-Cas9 *in Vivo* Gene Editing for Transthyretin Amyloidosis. *New England Journal of Medicine* **2021**, *385* (6). <https://doi.org/10.1056/nejmoa2107454>.
 13. Rusk, N. Human Immunity to Cas9. *Nature Methods* **2019**, *16* (4), 286–286. <https://doi.org/10.1038/s41592-019-0385-z>.
 14. Mehta, A.; Merkel, O. M. Immunogenicity of Cas9 Protein. *Journal of Pharmaceutical Sciences* **2020**, *109* (1), 62–67. <https://doi.org/10.1016/j.xphs.2019.10.003>.
 15. Frangoul, H.; Locatelli, F.; Sharma, A.; Bhatia, M.; Mapara, M.; Molinari, L.; Wall, D.; Liem, R. I.; Telfer, P.; Shah, A. J.; Cavazzana, M.; Selim Corbacioglu; Damiano Rondelli; Meisel, R.; Dedeken, L.; Lobitz, S.; Mariane de Montalembert; Steinberg, M. H.; Walters, M. C.; Eckrich, M. J. Exagamglogene Autotemcel for Severe Sickle Cell Disease. *New England Journal of Medicine/The New England Journal of Medicine* **2024**, *390* (18). <https://doi.org/10.1056/nejmoa2309676>.
 16. Diorio, C.; Shaughnessy, P.; Farhadfar, N.; Advani, A. S.; Weng, W.-K.; Yousoufian, H.; Chen, G.; Nguyen Brown, F.; Habtermariam, B.; Kumar, L.; Patel, J.; Liu, S. (Gary); Goyal, S.; Simon, A.; Minella, A. C.; Teachey, D. T. BEAM-201 for the Treatment of Relapsed And/or Refractory (R/R) T-Cell Acute Lymphoblastic Leukemia (T-ALL) or T-Cell Lymphoblastic Lymphoma (T-LL): Initial Data from the Phase (Ph) 1/2 Dose-Exploration, Dose-Expansion, Safety, and Efficacy Study of Multiplex Base-Edited Allogeneic Anti-CD7 CAR-T-Cells. *Blood* **2024**, *144* (Supplement 1), 4838–4838. <https://doi.org/10.1182/blood-2024-203546>.
 17. Canver, M. C.; Smith, E. C.; Sher, F.; Pinello, L.; Sanjana, N. E.; Shalem, O.; Chen, D. D.; Schupp, P. G.; Vinjamur, D. S.; Garcia, S. P.; Luc, S.; Kurita, R.; Nakamura, Y.; Fujiwara, Y.; Maeda, T.; Yuan, G.-C.; Zhang, F.; Orkin, S. H.; Bauer, D. E. BCL11A Enhancer Dissection by Cas9-Mediated *in Situ* Saturating Mutagenesis. *Nature* **2015**, *527* (7577), 192–197. <https://doi.org/10.1038/nature15521>.
 18. MedlinePlus. *Hemophilia*. medlineplus.gov. <https://medlineplus.gov/genetics/condition/hemophilia/>.
 19. Dale, D. C.; Makaryan, V. *rpE-Related* Neutropenia. *GeneReviews*® **2002**, No. 20301705.
 20. Da Costa, L.; Leblanc, T.; Mohandas, N. Diamond-Blackfan Anemia. *Blood* **2020**, *136* (11), 1262–1273. <https://doi.org/10.1182/blood.2019000947>.
 21. Hoban, M. D.; Lumaquin, D.; Kuo, C. Y.; Romero, Z.; Long, J.; Ho, M.; Young, C. S.; Mojadidi, M.; Fitz-Gibbon, S.; Cooper, A. R.; Lill, G. R.; Urbinati, F.; Campo-Fernandez, B.; Bjurstrom, C. F.; Pellegrini, M.; Hollis, R. P.; Kohn, D. B. CRISPR/Cas9-Mediated Correction of the Sickle Mutation in Human CD34⁺ Cells. *Molecular Therapy* **2016**, *24* (9), 1561–1569. <https://doi.org/10.1038/mt.2016.148>.
 22. Gupta, A. O.; Sharma, A.; Frangoul, H.; Dalal, J.; Kanter, J.; Alavi, A.; DiPersio, J.; Eapen, M.; Jaroscak, J. J.; Ayala, E.; Ziga, E. D.; Rifkin-Zenenberg, S.; Minella, A. C.; Chen, G.; Chen, Y.; Chockalingam, P. S.; Lin, L.; Joseney-Antoine, M.; Ianniello, L.; Gardner, B. Initial Results from the BEACON Clinical Study: A Phase 1/2 Study Evaluating the Safety and Efficacy of a Single Dose of Autologous CD34⁺ Base Edited Hematopoietic Stem Cells (BEAM-101) in Patients with Sickle Cell Disease with Severe Vaso-Occlusive Crises. *Blood* **2024**, *144* (Supplement 1), 513–513. <https://doi.org/10.1182/blood-2024-204888>.
 23. DeWitt, M. A.; Magis, W.; Bray, N. L.; Wang, T.; Berman, J. R.; Urbinati, F.; Heo, S.-J.; Mitros, T.; Munoz, D. P.; Boffelli, D.; Kohn, D. B.; Walters, M. C.; Carroll, D.; Martin, D. I. K.; Corn, J. E. Selection-Free Genome Editing of the Sickle Mutation in Human Adult Hematopoietic Stem/Progenitor Cells. *Science Translational Medicine* **2016**, *8* (360), 360ra134–360ra134. <https://doi.org/10.1126/scitranslmed.aaf9336>.
 24. Hanna, R.; Haydar Frangoul; McKinney, C.; Pineiro, L.; Mapara, M. Y.; Chang, K.-H.; Jaskolka, M. C.; Kim, K.; Rizk, M.; Olubunmi Afonja; Lawal, A.; Walters, M. C. S264: EDIT-301 SHOWS PROMISING PRELIMINARY SAFETY and EFFICACY RESULTS in the PHASE I/II CLINICAL TRIAL (RUBY) of PATIENTS with SEVERE SICKLE CELL DISEASE USING HIGHLY SPECIFIC and EFFICIENT ASCAS12A ENZYME. *HemaSphere* **2023**, *7* (S3), e05170e0e05170e0. <https://doi.org/10.1097/01.hs9.0000967968.05170.e0>.
 25. Eyquem, J.; Mansilla-Soto, J.; Giavridis, T.; van der Stegen, S. J. C.; Hamieh, M.; Cunanan, K. M.; Odak, A.; Gönen, M.; Sadelain, M. Targeting a CAR to the TRAC Locus with CRISPR/Cas9 Enhances Tumour Rejection. *Nature* **2017**, *543* (7643), 113–117. <https://doi.org/10.1038/nature21405>.
 26. Stadtmayer, E. A. CRISPR-Engineered T Cells in Patients with Refractory Cancer. *Science* **2020**, *367* (6481). <https://doi.org/10.1126/science.aba7365>.
 27. Research, C. for B. E. and CASGEVY. FDA 2023.
 28. Cui, T.; Li, B.; Li, W. NTLA-2001: Opening a New Era for Gene Therapy. *Life Med* **2022**, *1* (2), 49–51. <https://doi.org/10.1093/life-medi/lnac036>.
 29. Li, T.; Yang, Y.; Qi, H.; Cui, W.; Zhang, L.; Fu, X.; He, X.; Liu, M.; Li, P.; Yu, T. CRISPR/Cas9 Therapeutics: Progress and Prospects. *Signal Transduction and Targeted Therapy* **2023**, *8* (1), 1–23. <https://doi.org/10.1038/s41392-023-01309-7>.
 30. Charlesworth, C. T.; Deshpande, P. S.; Dever, D. P.; Camarena, J.; Lemgart, V. T.; Cromer, M. K.; Vakulskas, C. A.; Collingwood, M. A.; Zhang, L.; Bode, N. M.; Behlke, M. A.; Dejene, B.; Cieniewicz, B.; Romano, R.; Lesch, B. J.; Gomez-Ospina, N.; Mantri, S.; Pavel-Dinu, M.; Weinberg, K. I.; Porteus, M. H. Identification of Preexisting Adaptive Immunity to Cas9 Proteins in Humans. *Nature Medicine* **2019**, *25* (2), 249–254. <https://doi.org/10.1038/s41591-018-0326-x>.
 31. Li, A.; Tanner, M. R.; Lee, C. M.; Hurley, A. E.; De Giorgi, M.; Jarrett, K. E.; Davis, T. H.; Doerfler, A. M.; Bao, G.; Beeton, C.; Lagor, W. R. AAV-CRISPR Gene Editing Is Negated by Pre-Existing Immunity to Cas9. *Molecular Therapy* **2020**, *28* (6), 1432–1441. <https://doi.org/10.1016/j.ymthe.2020.04.017>.
 32. Laurent, M.; Geoffroy, M.; Pavani, G.; Guiraud, S. CRISPR-Based Gene Therapies: From Preclinical to Clinical Treatments. *Cells* **2024**, *13* (10), 800. <https://doi.org/10.3390/cells13100800>.
 33. Hryhorowicz, M.; Lipiński, D.; Zeyland, J.; Słomski, R. CRISPR/Cas9 Immune System as a Tool for Genome Engineering. *Archivum Immunologiae et Therapiae Experimentalis* **2017**, *65* (3), 233–240. <https://doi.org/10.1007/s00005-016-0427-5>.
 34. ClinicalTrials.gov. A Safety and Efficacy Study Evaluating CTX001 in Participants With Transfusion-Dependent β -Thalassemia. ClinicalTrials.gov. 2025. <https://clinicaltrials.gov/study/NCT03655678>.
 35. ClinicalTrials.gov. A Safety and Efficacy Study Evaluating CTX001 in Participants With Severe Sickle Cell Disease. ClinicalTrials.gov. 2025. <https://clinicaltrials.gov/study/NCT03745287>.

36. ClinicalTrials.gov. Long-Term Follow-Up of Participants Treated With CTX001. ClinicalTrials.gov. 2025. <https://clinicaltrials.gov/study/NCT04208529>.
37. ClinicalTrials.gov. A Study of EDIT-301 in Participants With Severe Sickle Cell Disease (RUBY). ClinicalTrials.gov. 2025. <https://clinicaltrials.gov/study/NCT04853576>.
38. ClinicalTrials.gov. A Study of EDIT-301 in Participants With Transfusion-Dependent β -Thalassemia (EdiTHAL). ClinicalTrials.gov. 2025. <https://clinicaltrials.gov/study/NCT05444894>.
39. ClinicalTrials.gov. A Study of BEAM-101 in Participants With Severe Sickle Cell Disease (BEACON). ClinicalTrials.gov. 2025. <https://clinicaltrials.gov/study/NCT05456880>.
40. ClinicalTrials.gov. A Study of BIVV003 in Participants With Severe Sickle Cell Disease (PRECIZN-1). ClinicalTrials.gov. 2025. <https://clinicaltrials.gov/study/NCT03653247>.
41. ClinicalTrials.gov. Long-Term Follow-Up Study of Participants Treated With BIVV003 or ST-400. ClinicalTrials.gov. 2025. <https://clinicaltrials.gov/study/NCT05145062>.
42. ClinicalTrials.gov. A Study of ST-400 in Participants With Transfusion-Dependent β -Thalassemia (THALES). ClinicalTrials.gov. 2023. <https://clinicaltrials.gov/study/NCT03432364>.
43. ClinicalTrials.gov. A Study to Evaluate SB-FIX in Participants With Hemophilia B. ClinicalTrials.gov. 2024. <https://clinicaltrials.gov/study/NCT02695160>.
44. ClinicalTrials.gov. A Study of OTQ923 in Participants With Sickle Cell Disease. ClinicalTrials.gov. 2026. <https://clinicaltrials.gov/study/NCT04443907>.
45. ClinicalTrials.gov. A Study of CTX001 in Pediatric Participants With Severe Sickle Cell Disease (CLIMB-151). ClinicalTrials.gov. 2025. <https://clinicaltrials.gov/study/NCT05329649>.
46. ClinicalTrials.gov. A Study of CTX001 in Participants With HbSC Disease. ClinicalTrials.gov. 2025. <https://clinicaltrials.gov/study/NCT05951205>.
47. ClinicalTrials.gov. A Study of CS-101 in Participants With Severe Sickle Cell Disease. ClinicalTrials.gov. 2026. <https://clinicaltrials.gov/study/NCT06565026>.
48. Shah, B. D.; Ghobadi, A.; Oluwale, O. O.; Logan, A. C.; Boissel, N.; Cassaday, R. D.; Leguay, T.; Bishop, M. R.; Topp, M. S.; Tzachanis, D.; O'Dwyer, K. M.; Arellano, M. L.; Lin, Y.; Baer, M. R.; Schiller, G. J.; Park, J. H.; Subklewe, M.; Abedi, M.; Minnema, M. C.; Wierda, W. G. KTE-X19 for Relapsed or Refractory Adult B-Cell Acute Lymphoblastic Leukaemia: Phase 2 Results of the Single-Arm, Open-Label, Multicentre ZUMA-3 Study. *The Lancet* **2021**, *398* (10299), 491–502. [https://doi.org/10.1016/S0140-6736\(21\)01222-8](https://doi.org/10.1016/S0140-6736(21)01222-8).
49. Overview | *Exagamglogene autotemcel for treating severe sickle cell disease in people 12 years and over* | Guidance | NICE. Nice.org.uk. <https://www.nice.org.uk/guidance/ta1044>.
50. National Academies of Sciences, E. *Human Genome Editing: Science, Ethics, and Governance*; National Academies Press, 2017.
51. *Current capabilities for human genome editing: report for the WHO expert advisory committee on developing global standards for governance and oversight of human genome editing*. www.who.int. <https://www.who.int/publications/i/item/WHO-SCI-RFH-2021.06-eng>.
52. Hickstein, D. D.; Calvo, K. R. The Spectrum of GATA2 Deficiency Syndrome. *Blood* **2023**, *141* (13), 1524–1532. <https://doi.org/10.1182/blood.2022017764>.

■ Author

Shriyans Katukuri, currently in the 11th-grade IB program at Interlake High School, is deeply devoted to understanding the processes behind mutations and their solutions. His

academic work centers on CRISPR and its growing range of applications, alongside the broader scientific principles in cancer biology.

Suitability of SAPARD Methods in the Post-War Recovery of Ukraine's Agricultural Economy

Sofiia Yanchuk

Queen Anne's School, Caversham, Berkshire, RG4 6DX, United Kingdom; s.m.yanchuk@gmail.com

Mentor: Georgios Aivaliotis

ABSTRACT: The Russo-Ukrainian war poses challenges for Ukraine's previously leading agricultural sector. In this paper, we aim to investigate how SAPARD (Special Accession Program for Agriculture and Rural Development) methods could benefit Ukraine's agricultural sector during the war. We provide a short literature review of Ukraine's needs in the agricultural sector recovery, and take Poland's SAPARD as an example of a successful agriculture modernization strategy during the pre-EU accession period. Then we use comparative analysis to evaluate the extent to which the same methods would be beneficial for Ukraine and which differences might exist in a similar program in Ukraine. The research proves that a SAPARD program would be partially beneficial for rebuilding Ukraine's agriculture, including the implementation of Measures 1, 2, and 3: Improving the Processing and Marketing of Agricultural and Fishery products, Investment in Agricultural holdings, and Development and Improvement of Rural Infrastructure. These Measures would increase agricultural productivity in Ukraine, restore transport links, encourage demining, and provide Agrocompanies with modernized machinery and technologies, while Measures 4, 5, 6, and 7 would mainly be helpful for Ukraine's EU integration. Additionally, we describe some differences in the organization of SAPARD Measures in Ukraine to achieve maximum efficiency.

KEYWORDS: Plant Sciences, Agriculture and Agronomy, Ecology, Agricultural Economics, War in Ukraine, SAPARD.

■ Introduction

The Russian-Ukrainian war, which started on the 24th of February 2022, has brought significant human losses and damage to all sectors of the economy. It also significantly affected the Agricultural sector, accounting for around 10% of Ukraine's GDP and providing 14% of employment in the pre-war period.⁸ According to the RDNA4 (2024) annual report, total agricultural damage is US\$11 billion, accounting for 6% of the total damage.¹³ On average, 200 kilograms of crops per hectare, including sunflower, wheat, barley, corn, and soybeans, were damaged by the war in 2022.¹¹ Given all the above, planning for the modernization, recovery, and reconstruction of Ukraine's agriculture in the post-war period is critical.

Several research papers and policy papers suggest improvements, while also emphasizing that the recovery of Ukraine will take place in parallel with EU integration.⁸ While searching for past project frameworks, researchers often turn to the experience of Poland, which is characterized by similar climate conditions and which has gone through a similar process of modernization in the early 2000s as part of SAPARD (Special Accession Program for Agriculture and Rural Development). Thus, they recommend considering SAPARD methods as part of Ukraine's post-war recovery.³

The Special Accession Program for Agriculture and Rural Development, SAPARD, was established in 1999 in 10 countries, including Poland, on their way to EU integration. With the main aim of preparing farmers and food processors for the strict EU veterinary and sanitary norms, it focused on agricultural modernization and rural development in Poland. The program's funds came from 2 sources: EU funds (75%) and

national co-financing (25%).¹⁰ The program's outcomes in the Lodz region included increased efficiency in food production, enhanced export capabilities, and improved quality of raw materials.²

In this paper, we aim to answer the question: What are the predicted effects of SAPARD methods on the Ukrainian recovery? For this aim, we use comparative analysis to examine how the potential implementation of each of the 7 SAPARD Measures in Ukraine's post-war recovery would benefit the restoration of the Agricultural sector of the economy. We provide a brief overview of the benefits and drawbacks of SAPARD, along with the primary methods used and the program's structure and organization. Then we compare Ukraine's current needs in the Agricultural Sector and use a comparative analysis to understand to what extent each Measure will work for Ukraine and which main aim each Measure will achieve. We also outline actions that could be taken to minimize the limitations of a SAPARD-like program in Ukraine, based on the Polish experience and the program's organization.

Additionally, we outline several crucial differences between the situations in Poland and Ukraine, which pose further limitations: the SAPARD program was part of Poland's transition to the EU. At the same time, Ukraine's key aim is post-war recovery.

The rest of the paper is organized as follows: In the next section, we include a short literature review highlighting the importance of agriculture in Ukraine's economy and the most commonly outlined needs and strategies for Ukraine's restoration, based on past papers. Next, we provide an overview of each of the 1-7 SAPARD Measures in Poland, including the

key benefits and limitations of their implementation towards overall economic performance in Poland and, specifically, in the Lodz region during the EU accession period. We also discuss the benefits and limitations of the organization and supervision of SAPARD performance in Poland. This section is followed by the comparative analysis, where we go through each of the Measures, compare their main aims and achievements with Ukraine's current needs, and identify possible drawbacks or differences in the outcomes. Finally, we use the available literature to discuss the sources of funding for a SAPARD-like program in Ukraine and outline possible limitations the program might face in Ukraine.

■ Methods

Background Information:

Over the 2001-2018 period, agri-food exports in Ukraine increased by 5.5 times, while imports rose by only 2.4 times in absolute terms, indicating that Ukraine was a net exporter (Figure 1).⁸ Ukraine became the world's leading sunflower oil exporter, with its main trading partners accounting for up to 45% of revenue: Asia (32%), Europe (32%), and the Middle East and North Africa (14%). Ukraine's significant contribution to global food security is feeding 332 million people. The share of the Agricultural sector in a country's GDP ranged from 6.5% to 14% during the 2001-2018 period. At the same time, the share of the population employed in the Agricultural sector was 15% in 2018, higher than in other European countries (Germany: 1%; France: 3%), according to the World Bank's open data.¹³ Grain is one of the most important crops produced in Ukraine. In the land structure, the highest yields were observed in the central part of Ukraine, with yields in other regions lower. This leads us to conclude that the war's impact and the occupation of 12% of mainly Eastern and Southern Ukraine significantly affected the production of the above-mentioned agricultural products.

At the start of the war in 2022, around 25% of rural households reduced or stopped agricultural production, mainly in the Southern and Eastern regions of Ukraine. At the beginning of the war, a significant decline in production was observed across all agricultural goods, with the most pronounced effects in vegetable and leguminous crops, which fell by 21.2% and 21.4%, respectively. It is important to mention not only the production of crops themselves, but also the amount of arable land available, which has declined due to mined areas that now cover up to 40% of Ukraine's territory. In addition to damage to production capacity, the author also highlights significant losses in the number of agricultural enterprises and equipment.⁶ The impact of the war on agricultural production was also analyzed using the AIRMAX model, which compared theoretical yields with actual yields in 2022 for the 6 most important crops in Ukraine: wheat, sunflower oil, rapeseed, barley, corn, and soy. The slightly different results from the above showed declines of 346 kg/ha for barley, 136 kg/ha for wheat, and 246 kg/ha for sunflower oil.¹¹ The total need for rebuilding Ukraine's Agriculture is \$11.2 billion as of the very beginning of 2025.¹³

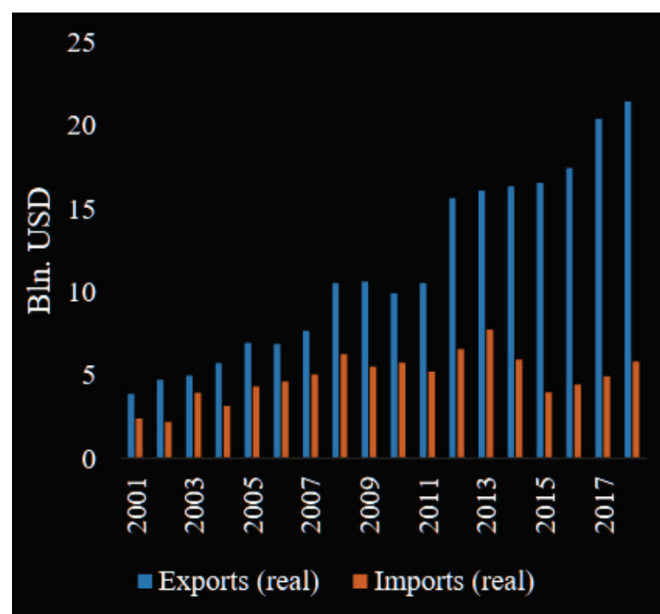


Figure 1: Real agri-food trade in Ukraine 2001-2018 in USD bln (in 2010 prices). It shows that Ukraine was the net exporter of agricultural production in the 2001-2017 period.⁸

The rebuilding of Ukraine's Agriculture should proceed through 4 stages, from urgent needs after the ceasefire to long-term EU integration over 5 years. At all stages, the attraction of foreign investment, modernization of agricultural equipment, restoration of export routes, specifically from the Black Sea, diversification, and retraining of the rural workforce are prioritized.¹ Attraction of the rural workforce and their training to identify landmines is highlighted as a key element in the post war rebuilding, in addition to the previously mentioned land-cleaning initiatives. Notably, the paper highlights the SAPARD program's successful experience in modernizing equipment and machinery in Central and Eastern Europe and the potential benefits of similar actions in Ukraine.³ Poland is cited as a case study of a country that followed a similar path in 2004. The SAPARD program is mentioned there as one that has primarily benefited the modernization of the agricultural sector. The research powerfully highlights the potential benefits for Ukraine from implementing a similar recovery strategy, with a specific focus on small-scale farms.¹²

Most of the researchers agree that the key initial step is to restore the functioning of the market at the maximum possible level, including demining and the alignment of the export routes with European standards, and to focus on restoring and modernizing the machinery and equipment, enhancing export routes, demining, and increasing productivity. After that, the focus should be on aligning Ukraine's Agriculture with EU norms, including sanitary and veterinary standards. For all stages, international support and funding are needed. The SAPARD program is mentioned several times as a successful example adopted by Central and Eastern European countries on the path to EU integration. We need to understand what the program methods entail, which of them led to the most significant improvements, and what might be used by Ukraine. Additionally, we should focus on the program's limitations and the differences between the SAPARD participants and

post-war Ukraine. For this research, we analyze the SAPARD program in Poland, given its similar post-Soviet past, climate conditions, crops produced, and relatively recent EU membership

What is SAPARD, its results, and limitations?:

The Special Accession Program for Agriculture and Rural Development, SAPARD, was established in 1999 in 10 countries, including Poland, on their way to EU integration. With 75% funding from the European Union (EUR 708.2 mln) and 25% (EUR 235.8 mln) from local co-financing, SAPARD in Poland, which accounted for 7% of GDP in 1995, focused on modernizing agriculture and rural development.

Before the implementation of SAPARD in Poland in 1884, 1,884 farms above 1ha were private, and 1.6 were public. The farms of 5-10 ha were the largest group in Poland, accounting for 24% of all farms. (Broyaka *et al.*). The organization of SAPARD in Poland was carried out by 2 main, very centralized organizations: ARMA (Agency for Restructuring and Modernization of Agriculture) and the Ministry of Agriculture and Rural Development (MARD), under the control of the European Commission. The process for farmers included confirming eligibility and submitting applications for specific Measures, meaning that individual farms could apply to all Measures or only to several of them. In case of success, 50% of the grants were distributed to farmers for the development of a specific Measure. The mid-term evaluation, monitoring, and control of the performance.⁷

In Poland, several measures of the SAPARD program were implemented.^{2,7}

Measure 1: Improving the Processing and Marketing of Agricultural and Fishery Products

Before the implementation of the SAPARD measures, Polish farms were characterized by low efficiency and obsolete machinery. Given the high unemployment level in Agriculture at this time, the products were of low quality and low value added. This means that the improvements in this Measure enabled higher efficiency in the processing of Agricultural and Fishery products, allowing farmers to process more raw materials into complete products. It included the installation and development of the machinery and the more efficient processing schemes. It then allowed farms to generate higher profits by increasing their value added, thereby directly improving the sector's profitability.

Measure 2: Investment in Agricultural holdings

Together with Measure 1, Measure 2 focused on generating value added and increasing profits. Unlike Measure 1, Measure 2 focused on product improvements and modernization, leading to the creation of new meat and milk production and increased environmental protection under this scheme. This and the previous Measure were key in driving the modernization and development of the agricultural holdings, which led to increased welfare for the farmers working in the industry.

Measure 3: Development and Improvement of Rural Infrastructure

This was one of the largest problems affecting the attractiveness and attracting foreign investment to Polish agriculture. The rapid development of this section of Economics in the late 1990s resulted in an imbalance in the allocation of infrastructure facilities. For example, only approximately 15% of the farms were connected to sewage systems, and the average age of the tractors was 16 years. The same problem was with the gas supply and the dumping sites. The SAPARD program thus solved this problem by directly providing these facilities. As a result, there is again greater worker and investor attractiveness towards rural areas. As a result, it began the Multiplier effect and brought significant income to Polish agriculture.

Measure 4: Diversification of Economic Activities in the Rural Area

This Measure was an important one for Polish Agricultural development, mainly due to the high level of unemployment and the obsolete machinery in use. Diversified economic activities created the potential for higher employment in the industry and increased farm profitability. The key improvements were achieved by financially supporting the farms and farm holders through direct investments and financing for tourist facilities. However, the Measure was delivered with a significant delay due to organizational issues, thereby limiting the results. Additionally, under the new activities, none were developed for environmental protection, indicating a lack of sustainability actions in Poland compared to other EU members. The Measure, which could have been strong, was limited by the organization and the regional government's lack of attention to it.

Measure 5: Agri-Environmental measure ups and Afforestation (was withdrawn from the program later and implemented under another program in 2004-2006)

Together with Measures 6 and 7, which were considered complementary Measures, less attention was paid to their development. This Measure was later canceled, so overall performance in this section was limited. Given that environmental protection was a marginal part of Measure 6, this suggests that the aim of increasing farm sustainability was the smallest among the other measures.

Measure 6: Vocational Training for Beneficiaries and Measure 7: Technical Assistance for Beneficiaries

These measures primarily focused on training in the use of new technologies and machinery to maintain high efficiency and increase farm profitability. Even though the Measures helped farmers gain basic knowledge of food processing, they focused almost entirely on veterinary, sanitary, and environmental measures, which were identified as important factors for EU accession. Thus, these Measures were complementary, meaning that their underdevelopment did not result in a decline in SAPARD benefits.

Analyzing the structure of the applications, around 85% of farmers in the Lodz region applied for diversification of the sector, even though this Measure accounted for only 10% of the available funds. The most significant part of the program focused on modernizing the agricultural holdings, supporting

farmers in buying agricultural machinery, upgrading agricultural production to meet EU standards, building roads, enhancing basic services, and promoting economic diversification. (Figure 2) A significant benefit of the program was the experience in using European Union funds for agriculture, which was one of Poland's first at the time. Participation in the program was mainly beneficial for large and medium-sized agrohholdings, the development and modernization of which later led Poland to become one of the major exporters of agricultural products in Europe, especially meat, while grain exports decreased. What is more, the program had a significant impact on the quality and sanitary adjustments to EU standards.²

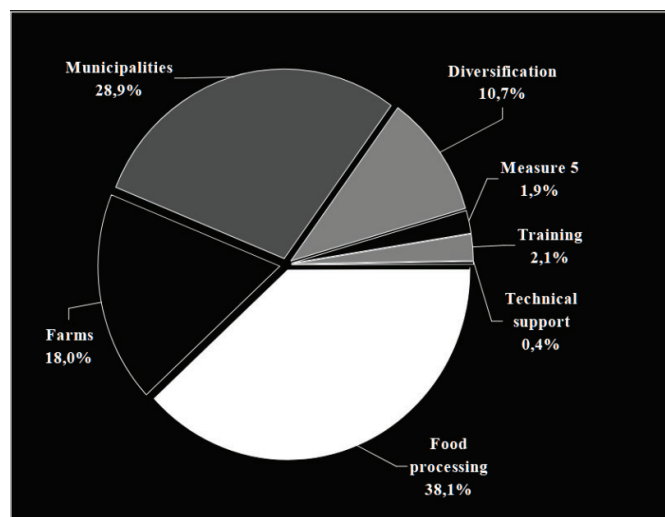


Figure 2: The planned structure of SAPARD expenditure in Poland. It shows that the greatest amount of funds was spent on the Food Processing (38.1%).²

At the same time, the application and organization model was imperfect, as the spatial differences in regional development were neglected and provincial authorities were excluded from decision-making and resource allocation. This means that the lowest control level was the regional one, which could not account for all the limitations and needs associated with SAPARD individual performance. These problems became visible as a result of SAPARD, which mainly benefited richer regions and large-scale agriholdings. Due to the lack of communication between regional and centralized authorities, the small- and medium-sized agrohholdings almost did not benefit from the program. Additionally, the program's limitation was a 50-year age restriction on farmers' applications, so anyone over this age was not eligible to apply. This means that only some of the farms could have benefited from the mentioned measures, whereas others' needs were neglected.⁷

In Poland, the increase in productivity on large-scale farms led to higher farm profitability and a stronger economy overall. The incomes of program beneficiaries increased by 20% on average. Thus, it led to economic growth, development, and modernization of agricultural holdings in Poland. The main benefits of SAPARD were associated with the renovation of obsolete machinery and rural infrastructure and the attraction of investments into the sector, as well as the production diversification, which allowed the increase in the value added

of the production, helping Poland to become one of the main exporters of Agricultural goods among the EU members. At the same time, the main aim of the program, which was rural development, was not fully achieved, as shown by measures such as unemployment and limited improvements in environmental initiatives. This was mainly due to problems with the organization and the underestimation of some of the Measures, including Measures 5, 6, and 7. These Measures were mainly responsible for the development and implementation of sanitary, veterinary, and ecological measures, and their underdevelopment set Poland partially behind the European standards. The organization and delivery of the program were also limited for smaller farms, older participants, and less wealthy regions, which made the extent of the program much smaller than expected at first.

Suitability of SAPARD in the rebuilding of Ukraine's Agriculture:

This research focuses on qualitative comparative policy analysis to understand the extent to which, and how, the SAPARD program's methods would benefit Ukraine's post-war agricultural recovery. To do this, we analyze the implementation process of each of the Measures in Poland, including their main successes and limitations. After this, we compare the outcomes in Poland with the situation in Ukraine and decide to what extent the strategies might be suitable. This will help us decide when, and if, the SAPARD methods should be implemented.

Measure 1: Improving the Processing and Marketing of Agricultural and Fishery products.

According to the papers discussed above, the most important agricultural products in Ukraine were wheat, sunflower oil, rapeseed, barley, corn, and soy. All of these have experienced significant declines due to the war and cannot be produced now at the same capacity as before the full-scale invasion. This leads us to the thought that the income generated from the exports of these products is not as high as before, and will not be at the same capacity level right after the war, because some territories, especially in the South and East of Ukraine, will require demining, which will take a significant amount of time.

In Poland, Measure 1 worked to increase the value added in the production of fishery and meat by expanding processing and marketing of agricultural goods, including the renovation of old techniques to increase production efficiency. As a result, Polish agricultural holdings became more competitive on the European Market and were able to meet European sanitary standards, while individuals gained higher profits.

Higher efficiency and value added in agricultural production would be a highly suitable way to recover Ukraine's Agriculture after the war. For example, wheat grown in the central part of Ukraine, which has experienced a minor impact of the war, might be processed into white flour. Corn might be processed into corn flakes or popcorn, thereby increasing its value. The improvements in the processing and marketing of agricultural goods in Ukraine will improve net trade, generating higher incomes within the economy and enabling the use of funds for demining the affected territories and the recovery of the

rural economy. Under this, higher employability in rural areas may contribute to labor reintegration. It is an important factor in the country's recovery, as the affected areas will gradually return to life.

Additionally, the wide range of production in the economy is a key element of EU integration for Ukraine. This means that the processing of agricultural goods and their compliance with EU sanitary standards will likely lead to faster integration in the long run.

Even though Measure 1 is well-suited to Ukraine and will lead to significant economic development, including greater profitability, post-war recovery, and EU accession, there may be some problems and limitations in implementing this Measure. First of all, Poland had functional, highly efficient logistics during the Measure implementation period, which Ukraine does not have. This will limit the extent to which agricultural goods can be exported. Additionally, the 40% land mining in Ukraine is a factor preventing the safety of the Measure's implementation. It will be dangerous to increase production in the rural regions straight after the war. Lastly, the great variety of goods produced was the main aim of Poland in this Measure. In contrast, in Ukraine, 30 years later, processing is already at a highly competitive level, so the impact of the Measure will not be as obvious as it was for the applicants in Poland.

Still, the limitations for the Measure implementation are mostly short-term, meaning that even if the exports of the goods will be more challenging in terms of logistics and land demining in the first stages, in the long term, it will increase the profitability of the individual holdings and improve the net trade, leading to the growth within the economy overall. The renovation of the obsolete and damaged machinery will lead to long-term value-added generation and an additional source of income, which can then be used to tackle other war-related damage.

Measure 2: Investment in Agricultural holdings.

In Poland, this scheme worked by investing in the functionality and productivity of farm equipment. Before implementation, Polish agriculture was characterized by obsolete machinery and inefficient use of technology, which set the economy back compared to other EU members. The Measure functioned by providing co-financed support for farms to purchase modern machinery. The success of the SAPARD Measure 2 thus depended on the large investments in purchasing new machinery. This helped the country modernize its production, thereby improving productivity and value added, which led to economic growth.

This Measure would be a powerful tool for returning the potential and productivity of Ukrainian farms. It might be done by replacing the damaged machinery in the South and East, and renovating the obsolete machinery in the West. This would help Ukrainian Agriculture meet EU requirements for modernizing techniques.

Even though the Measure is highly suitable for both economic recovery and EU integration, there may still be limitations to its development. First of all, the mined territories would not allow control over the functionality of this Measure

across all territories. Secondly, it is important to ensure that the applications have different aims, whether restoration or renovation. Thus, the Measure should include different application methodologies for Western Ukraine and South-Eastern Ukraine. The greatest amount of funds should be allocated to restoring the damaged machinery after the war. In contrast, a smaller amount should be allocated to renovating obsolete machinery in the West. Also, resources should first be allocated to post-war recovery and only then to renovation and modernization.

Still, this Measure is highly suitable for the recovery of Ukraine's Agriculture after the war and for EU accession. It is important to note that this Measure will only work efficiently if Measures 6 and 7 are not limited as they were in Poland during SAPARD. In this case, the use of the new machinery will be the most efficient, and the highest profits could be generated. However, overall, the modernization and restoration of the Ukrainian farms at a high level can be achieved through Measure 2. If the government controls the division of funding between restoration and modernization, the program will be the most successful. This is one of the most important SAPARD measures that should be implemented in Ukraine as well.

Measure 3: Development and Improvement of Rural Infrastructure.

The development and improvement of rural infrastructure helped Poland build the sewage systems and networks needed for the more efficient operation of the holdings. The direct results were the sector's greater attractiveness to investors and the higher profitability of the holdings. Additionally, production became easier and more environmentally friendly, allowing Poland to meet European standards and become one of the largest exporters of agricultural products in the European Union.

The main reasons for the quick success of this Measure of SAPARD were the stability of the organizational institutions. What is different from the Ukrainian situation is that Poland did not experience any infrastructure damage, so the main goal of Measure 3 was the modernization of the existing system rather than the repair of damaged infrastructure. The main actions done by Measure 3 were in the installation of the new system.

Measure 3 is one of the key factors that should be implemented for the redevelopment of Ukraine. As seen in the previous sources, transporting agricultural products in Ukraine is often difficult due to poor road conditions, and, as CEMA shows, railway services are unsuitable for export due to war-related destruction. Additionally, the war's damage includes infrastructure. With improved infrastructure, farmers would be able to expand their farms and attract workers, which is important given that over 6 million Ukrainians emigrated after the beginning of the war, and 600,000 are now part of the national army. In the long term, this will improve logistics services and networks, allowing higher profitability and efficiency in imports, while the quality of agricultural production will also improve.

However, the key difference and potential problem with the implementation of the Measure is that most of the rural infrastructure is either partially obsolete in Western Ukraine or totally damaged in the South and East of Ukraine. So the implementation of Measure 3 will involve both the installation of the new networks and the removal of the damaged ones, in addition to their modernization, as in Poland. This means there may be a higher cost to the organization. For the Measure to be effective, it is also important that the government and regulatory bodies ensure that the highest amounts of financial resources and, thus, the highest number of applications come from Eastern and Southern Ukraine. In contrast, applications from Western regions are mostly focused on modernization and are less price-sensitive. If this is not controlled, it might lead to inefficient resource allocation, where less damaged territories use a higher percentage of the funds. In this case, the situation might be similar to the one that happened in Poland, where only the wealthier farms achieved the greatest success.

To make the control and distribution of these main roles as effective as possible, there might be a plan to divide this Measure into 2 submeasures, each covering a different part of Ukraine and with different levels of funding. In this case, there is a high probability of the development and restoration of logistics services, transport networks, and rural systems, which will restore the potential of Ukraine as the biggest exporter of wheat and sunflower oil before the beginning of the war.

Measure 4: Diversification of Economic Activities in the Rural Area.

In Poland, the diversification of economic activities meant direct financing of the farming industry to expand its range of activities. The Measure aimed to modernize older farming techniques and promote the farming industry to youth to increase their engagement in the sector and their employment in it. The Measure was important for Poland because the agricultural sector was underdeveloped, and much of the land was unused in many areas. It was important as well, given the prevalence of the elderly in the agricultural sector and the shortage of young, well-educated professionals.

However, if you compare this Measure with the current situation in Ukraine, a large share of economic activity is conducted in rural areas. Additionally, arable land accounts for 56.8%, and most farms' productivity is relatively high.¹³ There is already an established "status" for Ukraine on the international market as an exporter of wheat, oil, and corn. Ukraine's market share in international agricultural trade was very high, and major agroholdings, such as Kernel and MXII, deliver a wide range of activities.

From this analysis, we understand that this Measure will not have a significant impact on the restoration of the agricultural sector. This is because of the already high productivity before the war, the relatively well-developed land, and the production on it. The implementation of this Measure will require substantial management across different governance layers and a substantial financial commitment. However, the consequences will only be positive for the small farms, which will be able to maximize their profit. This would benefit the development of

the economy and the growth of the sector. However, the overall state of the sector is not as underdeveloped as it was in Poland before the EU accession. For Poland, it was the first step toward becoming a global exporter of agricultural production, whereas Ukraine had already taken this step before the 2022 war.

As part of the EU accession, Ukraine would benefit from the development of sanitary and veterinary standards for smaller and medium-sized holdings. In contrast, for war recovery, logistics and demining should be carried out first, with additional focus on attracting residents back to the rural areas and on their ability to produce goods using the renovated machinery and new techniques. Neither of these needs is outlined by Measure 4 of SAPARD in Poland, so the Measure will not be highly beneficial for Ukraine in the post-war period. However, this Measure might be done in the 5+ year perspective for the targeted areas with underdeveloped holdings in specific regions, which will require it most. However, this will probably be part of the internal, internally financed project undertaken by the Ukrainian government to gain even higher market share on the global market. This Measure is therefore not suitable for Ukraine's current needs in the post-war recovery and EU integration.

Measure 5: Agri-Environmental measure ups and Afforestation.

In Poland, this Measure was introduced to address environmental development in the area and to align with EU standards. However, much less attention and fewer funds were devoted to this Measure, which limited its overall performance. Later on, the withdrawal of Measure 5 from the SAPARD program highlighted that the main achievement of SAPARD was in the region's economic development, rather than alignment with EU veterinary and sanitary standards or environmental sustainability.

As part of EU integration, several sources from the literature review mentioned above recommend implementing this Measure in Ukraine. Also, in the case of Ukraine, it is likely associated with land demining after the war and the resolution of soil degradation problems, as currently up to 40% of the arable land in Ukraine is classified as contaminated by land mines. After the war, contamination of a large area was observed, and it is essential to address this problem. The Measure should be highly suitable for dealing with pollution and land management in Ukraine. It is also important to ensure these methods are implemented across different-scale farms, as most large agroholdings have the appropriate environmental conditions, while the smaller ones do not.⁸

In terms of afforestation, Ukraine has lost 12% of its forest cover over the last 24 years. To some extent, deforestation is a problem in the Eastern regions of Ukraine, where fires have destroyed significant areas. Still, the problem of deforestation is mainly associated with the Kyiv, Rivne, and Zhytomyr regions, which were not significantly affected by the war (Global Forest Watch).

Overall, this Measure has a long-term benefit for Ukraine when it focuses on alignment with the EU and environmental sustainability. However, this action will likely not be taken in

the first year after the war ends, as the main aims will be in the restoration of the logistic paths, land demining, and restoration of the damaged machinery to increase the productivity of farms and to begin to generate the profits, which can be later used for the co-financing of these environmental Measures. So Measure 5 will likely become the priority only in the long-term perspective of the recovery in 2+ years after the end of the

Measure 6: Vocational Training for beneficiaries and Measure 7: Technical Assistance for beneficiaries.

This is an important way of integrating Ukraine's Agriculture into the EU. Under the successful implementation of Measures 2 and 3, the great majority of the newly generated workforce would need to use the infrastructure and technologies. The implementation of vocational training would improve the productivity of the Agricultural holdings and generate more income. As a result, small-scale farms could turn into modern ones.

However, in Poland, the implementation included training individuals to use the new machinery rather than addressing the steps contributing to sanitary, veterinary, and environmental standards. This means that if these Measures are planned for the Ukrainian post-war recovery, it would be beneficial, though with a slightly shifting focus on EU integration. As there are no significant problems with the use of the machinery, but with environmental issues and development, the attraction of professionals from abroad would help. Additionally, the training may include topics such as land demining techniques, logistics, and the acknowledgment of new trading partners. Lastly, even though these Measures are highly suitable and useful for the functioning of Ukrainian agriculture after the war, their implementation will not be a priority after the end of the war, when the restoration of functionality and the survival of farms will be prioritized. They might be working for EU integration in the long-term 3+ year plan, but not as an immediate action.

Organization:

There should also be changes to the organization of SAPARD in Ukraine that take into account the challenges Poland faced previously. First of all, SAPARD control in Poland was mainly managed at the national level, with regional committees participating in decision-making. This put pressure on the controlling organizations, such as ARMA and MARD, as they needed to monitor performance and allocate funds to individual farms across the country. For this reason, the needs of many smaller agroholdings were neglected, and most of the funds were distributed to wealthier regions and larger farms.

In Ukraine, such a program must not be focused on the wealth of the farms but rather on their needs and destruction, which happened to the rural communities during the period of war. For this reason, a more diversified and less concentrated scheme of control and resource allocation should take place in Ukraine. This might involve the national controlling organizations sending limited resources to different regions of Ukraine, with higher funding for regions with the greatest damage to Agriculture. In this situation, the southern and eastern regions of Ukraine would benefit the most. Next, the regional govern-

ment should allocate resources within the region, with district organizations collecting information on the financial needs of individual farms and passing it to the regional government. In this case, the more complex system with more sub-levels of resource allocation will reflect the needs of individual farms based on their level of damage, not their wealth.

In terms of application strategies, it makes sense to retain the same scheme as in Poland, where individual farms can apply for several or all of the Measures. However, it makes sense to limit the number of applications from Western Ukraine, especially for the least important Measures in the restoration of

Agriculture, such as Measures 6 or 7. The relatively wealthy and less damaged Western part of Ukraine can organize Vocational and Technical training itself without additional funding from abroad. It also makes sense to split some of the Measures into 2 parts to better meet the needs of specific regions. For example, Measure 3: Development and Improvement of Rural Infrastructure can be divided into 2 subsections: development and improvement. The development section should focus only on the regions with agricultural damage, primarily the machinery. Whereas the Improvement Section can be mostly concentrated on the Western parts of Ukraine and the less damaged regions. The Improvement section will not create any new machinery or equipment, but will help modernize the old and inefficient techniques currently in use. This division will help ensure that the modernization of the Agricultural holdings is carried out across all territories. However, the focus is chosen according to the damage to the territories due to the war. In this case, Ukraine will meet 2 of the most important aims of EU integration and the post-war restoration.

Sources for funding in Ukraine:

It is important to understand that if the SAPARD program were used in Ukraine, it would require funding from abroad, as Ukraine is currently unable to cover its rebuilding needs on its own due to the war. Additionally, 75% of SAPARD funding was provided by the EU, so we should also consider it a source for Ukraine.

Recovery, rebuilding, and reconstruction of Ukraine will be largely integrated into the EU accession process.⁹ This means that a significant source of donations would presumably come from the EU, the World Bank, and the EBRD to ensure the formation of a competitive, diversified economy, and these investments for international companies are the most vital to kick-start reconstruction. Additionally, it is important to understand the differences between the situation in Poland and in Ukraine, and to consider that the current financial support provided to Ukraine during wartime is much higher than that provided to Poland during the pre-EU accession period. The organizations mentioned below, and the funding opportunities, illustrate that it is easier to implement the SAPARD methods in Ukraine.

There has been and is significant funding from international organizations that could be used to finance a SAPARD-like program. As part of its 2023 response plan for Ukraine, FAO had budgeted \$205 million to help rural households and small farmers living in the frontline areas. As of December 2024, the

FAO, in partnership with the EU, has provided grants to 237 agricultural producers for the purchase of equipment, green technologies, and infrastructure improvements, totaling USD 2.8 million. At the same time, USAID and the Ministry of Agrarian Policy and Food of Ukraine supplied fertilizers to farms in Ukraine. This is followed by USAID's provision of sunflower seeds for 42,000 hectares of land. The donations are crucial for Ukraine's Agriculture in restoring farm functionality and maintaining them in the first period after the war. The European Bank for Reconstruction and Development (EBRD) has also provided significant assistance to Ukrainian Agriculture and is expected to increase investments to \$3-5 billion. In addition to the global organizations above, there are specific organizations that primarily focus on supporting small agricultural holdings. IFAD (International Fund for Agricultural Development) is an institution that aims to help smallholders affected by the current conflict in Ukraine.^{4,5}

■ Results and Discussion

Results:

The comparative policy analysis helps to understand that SAPARD methods would generally be suitable for the recovery of Ukraine's Agriculture after the war. This is proved by the high correlation between the proposed reconstruction pathways for the Agricultural sector of Ukraine and the methods performed by the SAPARD program in Poland during the pre-EU accession period. Specifically, the most important measures are Measures 1, 2, and 3. They would create workplaces in the agricultural sector, generate profits, and encourage further investment. These measures follow the stated needs of Ukrainian Agriculture, as identified in previous research, namely the restoration and redevelopment of transport infrastructure, the modernization of agricultural equipment, land demining, and increased land productivity, which will lead to intensified trade opportunities, the attraction of a new workforce, and the generation of value added in agricultural production. Measures 4, 5, 6, and 7 would help with EU integration; however, the Environmental measures in Measure 5 can be used in the first stages of the post-war recovery, mainly as part of land demining. These Measures should definitely take place in the development of Ukraine's Agriculture, but at later stages, and not right after the ceasefire. The information on the main benefits and limitations of the SAPARD Measures in Ukraine is summarized in Table 1, while the overall results and findings are in Table 2. Additionally, the research helps to understand the potential sources for the future funding of the SAPARD-like program in Ukraine, which will include FAO, EBRD, and IFAD funds. This proves the idea that a similar program could be implemented in Ukraine with a relatively smaller contribution from the national side and higher international support, compared to Poland.

It is also important for Ukraine to have a slightly different organizational system for the program's success. It is beneficial for the district government to distribute the funds and collect the applications under the guidance of the regional and, later, the national government. This will ensure even opportunities

for larger and smaller farms, but the national government will still be able to direct most of the funds to the most destroyed territories of Ukraine.

Table 1: Main Benefits and Limitations of each of the SAPARD Measures implementation in Ukraine. The table shows that most of the Measures will result in higher productivity and modernization of agricultural production.

Measure	Main Benefit	Main Limitation
Measure 1: Improving the Processing and Marketing of Agricultural and Fishery products.	Increased range of production, increased value-added, and profits could be the sources for land demining	Processing of the Agricultural goods is already on the high level; land mining can limit the positive effect
Measure 2: Investment in Agricultural Holdings	New machinery purchased, old machinery renovated; increased productivity and profitability of the farms, including small-scale holdings	Investments can be used inefficiently due to the high land mining. The needs of the Southern and Eastern Ukraine in development differ from the needs of Western Ukraine in the improvement of the rural infrastructure; profound control and resource distribution are needed
Measure 3: Development and Improvement of Rural Infrastructure	Restoration of the logistic routes and the trade opportunities in the rural areas; attraction of the new workforce; growing trade and higher incomes of the residents	The needs of the Southern and Eastern Ukraine in development differ from the needs of Western Ukraine in the improvement of the rural infrastructure; profound control and resource distribution are needed
Measure 4: Diversification of Economic Activities in the Rural Area	Even higher incomes of the farms	Ukraine already has well-developed and diversified agricultural activities that can be financed internally
Measure 5: Agri-Environmental Measures and Afforestation	Land demining, decreasing level of soil pollution	Does not increase the productivity/profitability of the farms; only suitable for the long-term aims, not the urgent post-war needs.
Measure 6: Vocational Training for beneficiaries and Measure 7: Technical Assistance for beneficiaries.	Better usage of the machinery, higher productivity of agroholdings	Not the most necessary for the post-war reconstruction, EU integration might be limited due to the limited funds.

Table 2: Suitability of SAPARD methods for the recovery of Ukraine's Agriculture. The table shows that Measures 1, 2, and 3 are the most suitable for replication in Ukraine.

Measure of SAPARD	Suitability for Ukraine's Agriculture post-war recovery?	Any other potential benefits of applying this Measure?
Measure 1	High	EU integration
Measure 2	Critical	EU integration
Measure 3	Critical	EU integration
Measure 4	Low	EU integration
Measure 5	Moderate (only for land demining)	EU integration
Measure 6	Low	EU integration
Measure 7	Low	EU integration

Discussion:

After comparing the SAPARD program methods with Ukraine's current agricultural needs, we understand that this strategy will be most useful for Ukraine's post-war recovery, with a simultaneous focus on EU integration, thereby combining 2 ideas into 1 process. Moreover, Ukraine is currently receiving broad support from international organizations (the EU) and private companies (Bayer). It is expected to receive post-war funding for recovery, which would enable the implementation of SAPARD program methods even with smaller investments from national companies. However, there are several things to consider.

First of all, the reconstruction and recovery of Ukraine's Agriculture is a long-term process that requires long-term strategies. The SAPARD program's methods focus on improving the state of Agriculture, not rebuilding it after complete deterioration in some regions of Ukraine. Additionally, agriculture is vital in Ukraine, but it is still less important than human safety. So, the program methods could only occur af-

ter the water and food supplies are restored to the unoccupied regions, the establishment of a network and communication, and the restoration of law and order in Ukraine. Then the first Agriculture-related activity would be land demining, followed by the restoration of full export capabilities through the Black Sea port and railway transportation. Only after this can the methods be implemented, at a minimum, not earlier than 1 year after the ceasefire.

Another point is the drawbacks of SAPARD, which should be considered in Ukraine. It is stated that the largest agrohholdings received the most significant benefits, while small and medium ones received less support. It is crucial not to let this happen in Ukraine. During the war, the larger agrohholdings adapted relatively easily to the circumstances; thus, even with the previously mentioned changes to the organization of the SAPARD Measures in Ukraine, it might still be hard to determine which beneficiaries will benefit the most. In the worst case, it might lead to inequality among rural households across different regions of Ukraine.

■ Conclusion

In this research, we examined whether the SAPARD program methods, previously implemented in Poland during its EU integration in the early 2000s, would be partially beneficial for Ukraine in its post-war recovery. After reviewing the previous research on the primary post-war needs of Ukraine's agriculture and the effectiveness of the SAPARD program in Poland, we conducted a comparative study of the possible implementation of 7 different Measures used by SAPARD into Ukraine's agricultural economy.

The results show that 3 Measures: Measure 1: Improving the Processing and Marketing of Agricultural and Fishery products, Measure 2: Investment in Agricultural holdings, Measure 3: Development and Improvement of Rural Infrastructure would be beneficial for the improvement of the productivity of the farms, the modernization of the equipment and implementation of the technologies, land demining, and rebuilding the transportation links. Measures 4, 5, 6, and 7 are primarily associated with Ukraine's integration into the EU and can only be used in the later stages of Ukraine's rebuilding. Additionally, for the program to be beneficial in Ukraine, the organizational differences, including less centralized governance and a focus on the most damaged regions, should be implemented.

Although the program would have some important differences from the SAPARD in Poland, including the timing and priority of several measures, focus on the regions with the most significant need for implementation, and the equality of all of the beneficiaries, we suggest that the SAPARD program methods would be a valuable and helpful addition to rebuilding Ukraine's Agriculture after the war. Future research might include evaluating the effectiveness and urgency of SAPARD methods across regions of Ukraine. Additionally, reliable financial sources should be identified, and a detailed implementation plan for each Measure should be developed. The pilot programs might be beneficial for further testing.

■ Acknowledgments

We express our specific gratitude to Professor Georgios Aivaliotis and mentor Daryn Howland for supporting us throughout the research process, teaching, and assisting us with this work. We are also thankful to the IRIS research team for the additional support throughout.

■ References

1. Broyaka, A.; Dankevych, V.; Dodd, E.; Welsh, C. *Blueprint for an Agricultural Recovery Plan for Ukraine*; Center for Strategic and International Studies: Washington, DC, 2025. <https://www.jstor.org/stable/resrep70828> (accessed May 17, 2026).
2. Bryła, P. *The Role of SAPARD in Stimulating Entrepreneurship among Food-Processing Firms in the Lodz Region*, 2005. https://www.researchgate.net/publication/23510081_The_Impact_of_SAPARD_on_the_Behaviour_of_Farms_and_Food-Processing_Enterprises_in_the_Lodz_Region (accessed May 17, 2026).
3. CEMA. *Rebuilding Ukraine's Agriculture after the Full-Scale War: A Humanitarian and Economic Imperative*; European Agricultural Machinery Association: Brussels, 2023. https://www.cema-agri.org/images/publications/position-papers/CEMA-Rebuilding_Ukraines_Agriculture_after_the_War-Policy_Paper-2023-08.pdf (accessed May 17, 2026).
4. Gagalyuk, T. *Financing of Resilience and Reconstruction of the Agricultural Sector of Ukraine by International Financial Institutions and Development Agencies*, 2024. https://www.apd-ukraine.de/fileadmin/user_upload/APD_Taras_Gagalyuk_ENG.pdf (accessed May 17, 2026).
5. Kalina, I.; Shkolenko, O.; Shulyar, N.; Kosmidailo, I.; Palii, S.; Panchenko, O. Directions of Financing the Needs of Rebuilding the Agricultural Sector of Ukraine. *Finansy, Kredyt, Działalność* **2024**. <https://fkd.net.ua/index.php/fkd/article/view/4301> (accessed May 17, 2026).
6. Kotytkova, O.; Babych, M.; Nadvynychnyy, S.; Pohorielova, O. Agriculture in Ukraine: Regional and Global Impacts of the War. *European Countryside* **2025**. <https://doi.org/10.2478/euco-2025-0011>.
7. Matteus, D.; Hazners, J.; Kovacs, K.; Stoian, M.; Georgieva, M.; Cerna, P.; Grosse, T. *SAPARD Review in Bulgaria, Czech Republic, Estonia, Hungary, Latvia, Poland and Romania: Impact Analysis of Agriculture and Rural Development*, 2005. <https://cps.ceu.edu/publications/joint-publications/sapard-review> (accessed May 17, 2026).
8. Nivievskyi, O.; Martyshev, P.; Kvasha, S. Agricultural Policy in Ukraine. *arXiv* **2023**, arXiv:2305.01478. <https://doi.org/10.48550/arXiv.2305.01478>.
9. Nivievskyi, O.; von Cramon-Taubadel, S. *Rebuilding Ukraine: The Agricultural Perspective*, 2023. <https://www.ifo.de/DocDL/econpol-forum-2023-2-cramon-taubadel-nivievskyi-agricultural-perspective.pdf> (accessed May 17, 2026).
10. Oczkowska, M.; Myck, M. *FREE Policy Network Brief Series #AcademicsStandWithUkraine*, 2025. https://freepolicybriefs.org/wp-content/uploads/2025/02/20250220_boxes-komprimerad.pdf (accessed May 17, 2026).
11. Petrenko, O. *The Impact of the Full-Scale War on Agricultural Performance in Ukraine*, 2025. https://kse.ua/wp-content/uploads/2025/06/Petrenko-Oleksandr-BE25_-THE-IMPACT-OF-THE-FULL-SCALE-WAR-ON-AGRICULTURAL-PERFORMANCE-IN-UKRAINE.pdf (accessed May 17, 2026).
12. Warsaw Enterprise Institute. *Ukrainian Path to the European Union: Polish Experience – Agricultural Sector*, 2024. <https://wei.org.pl/wp-content/uploads/2024/11/Ukrainian-path-to-the-Europe->

an-Union.-Polish-experience.-Agricultural-sector.pdf (accessed May 17, 2026).

13. World Bank; Government of Ukraine; European Union; United Nations. *Ukraine – Fourth Rapid Damage and Needs Assessment (RDNA4): February 2022 – December 2024*; World Bank Group: Washington, DC, 2025. <http://documents.worldbank.org/curated/en/099022025114040022> (accessed May 17, 2026).

■ Author

Sofia Yanchuk is a Year 12 student at Queen Anne's School who decided to conduct the research to contribute to Ukraine's post-war recovery. Sofia is interested in Economics and Geography, a participant in the International Geography Olympiad. She is a dancer, a leader of the school government, and a volunteer.

Mathematical and Machine Learning-Based Modeling of Plant Recovery Dynamics Under Environmental Stress

Yenula Wickramasinghe

Bridgeland High School, 10707 Mason Rd, Cypress, TX 77433, USA; yenulawickramasinghe@gmail.com

Mentor: Rayvan Watson

ABSTRACT: Studying plant recovery from environmental stress and how recovery dynamics differ by stress type is important yet remains insufficiently understood. This study aims to model plant recovery dynamics under water and light stress, and to predict the growth of the plant under the stress. Using three mathematical models, recovery rates of the under-stress plants are modeled and compared for accuracy. Random Forest model is used to predict the growth of the plant exposed to stress. According to the experimental findings, light-stress recovery is best described by exponential behavior, agreeing with the rapid, reversible physiological adjustment, whereas water-stress recovery follows power-law scaling, indicative of prolonged, scale-dependent recovery. A Random Forest model further showed that it can be used effectively to forecast the future growth of the plant under stress. Overall, these findings indicate the importance of time-resolved mathematical modeling for understanding plant resilience and developing predictive tools for adaptive agricultural management.

KEYWORDS: Plant Recovery Modeling, Environmental Stress in Plants, Nonlinear Regression, Machine Learning in Agriculture, Predictive Modeling.

■ Introduction

Plants are continuously exposed to changes in environmental conditions such as water limitations, reduced light, extreme temperatures, and nutrient deficiencies. Compared to the studies on the physiological impacts of environmental stress, relatively less attention has been given to understanding the dynamics of plant recovery following stress removal. Most of the past studies focused on the damage induced by stress at a single point instead of characterizing how recovery progresses over time.¹ However, studying recovery kinetics will reveal how quickly plants return to the baseline, providing deeper insight into stress resilience.

The impact of the water and light stress on plants takes place in distinct physiological pathways. Water limitation disrupts plant functions, such as cellular turgor pressure, stomatal conductance, photosynthetic efficiency, and metabolic balance.² Light-stress impacts plants primarily with photosystem performance and chlorophyll dynamics.¹ The above two stresses operate through different mechanisms and are different in their recovery trajectory and structural patterns.

Recent research indicates that these response patterns to stress depend on the plant's exposure to prior stress, and it forms a phenomenon known as "stress memory."³ This suggests that recovery is not merely passive but a time-dependent biological process. Furthermore, environmental stress responses involve multi-scale interaction across cellular and whole plant levels, suggesting the entire recovery dynamics may follow non-linear behavior.⁴

With the understanding that the plant stress recovery is a non-linear time-dependent process, it is interesting to know how the recovery progresses: Does it occur at a constant rate? Does it follow an exponential path toward stability? Math-

ematical modeling of the recovery process will provide a framework to formally characterize these dynamics. Linear models describe a constant rate of recovery, exponential models capture acceleration toward stability, while power-law models allow exploration of scale-dependent behavior often used in biological growth systems.⁴

Recent advances in data-driven modeling, such as machine learning models, extend this framework. Machine learning algorithms allow recovery patterns to be treated and forecasted without enforcing predetermined parametric forms. By incorporating biological characteristics of the plants, such as plant height, leaf angle, and leaf color, into predictive algorithms, these machine learning algorithms allow us to estimate future recovery states based on present measurements. This integration of predictive models offers a pathway toward precise agricultural applications, easing the process of irrigation scheduling and adaptive management strategies.

Although many previous studies have examined how environmental stress damages plants, most have focused mainly on the immediate effects of stress rather than on how plants recover over time once the stress is removed. Much less is known about how recovery patterns differ between different types of stress, such as water and light limitation, or whether these recovery processes are best explained by different mathematical models. In addition, only a limited number of studies have combined traditional biological modeling with machine learning to both analyze and predict plant recovery. To address this gap, the present study investigates how plants recover over time from water and light stress, compares multiple mathematical models to describe these recovery patterns, and incorporates machine learning methods to improve prediction and understanding of plant resilience.

This manuscript investigates the time-resolved recovery dynamics of plants exposed to water and light stress. Combining controlled experimentation, nonlinear regression, and machine learning modeling, this work aims to (1) quantify the recovery trajectories of the plant due to the two types of stresses, (2) assess the appropriateness of linear, exponential, and power-law models for describing recovery kinetics, and (3) study the feasibility of forecasting the next-day recovery states.

■ Methods

Experimental Setup:

A total of 30 snow pea plants were grown under identical indoor conditions for two weeks prior to stress application. Plants were then randomly assigned into three groups: water-stress ($n = 10$), light-stress ($n = 10$), and standard-condition group ($n = 10$). For the first group, water stress was applied, and for the second, light stress was applied. The third group was considered the control group. Water stress was applied through controlled reduction of irrigation, while light stress was introduced by limiting light exposure for a defined period. For the control plants, standard conditions were maintained under consistent irrigation and light schedules.

During recovery, plant height, leaf angle, turgor, and leaf color were measured repeatedly at consistent time intervals to capture time-resolved responses. Plant height was documented using a metric ruler from the soil surface to the highest point of the stem, and leaf angle was measured using a handheld protractor relative to the main stem. Measurements were performed by the same observer using identical handling pressure and assessment criteria at each time point to minimize subjective variability. Leaf color and turgor were evaluated using predefined visual assessment scales under consistent indoor lighting conditions.

Measurements were performed at similar times of day using consistent procedures to ensure consistency across plants and time points. Qualitative physiological observations, such as leaf color and turgor, were converted into standardized numerical scores using predefined rating scales.

Recovery function $R(t)$:

We normalize the collected data of height, leaf angle, turgor, and leaf color and then compute the recovery index, $R(t)$ as follows.

Then, for each observation under light or water stress, the measurements were normalized:

$$\text{normalized metric} = \frac{\text{metric}(t)}{\text{baseline mean}} \quad (1)$$

A composite recovery index was then defined:

$$R(t) = \frac{1}{4}(\tilde{H}(t) + \tilde{A}(t) + \tilde{T}(t) + \tilde{C}(t)) \quad (2)$$

where the $\tilde{H}, \tilde{A}, \tilde{T}, \tilde{C}$ variables represent normalized height, leaf angle, turgor, and leaf color.

Modeling Approach:

Plant recovery following environmental stress was modeled as a time-dependent process, $R(t)$, where time $t \geq 0$. Although recovery is continuous, measurements were obtained

at discrete, evenly spaced time points: $R(t_i)$, $i = 1, 2, \dots, n$, where $\Delta t = t_{i+1} - t_i$ is the sampling interval.

Linear Model:

A baseline linear model was defined as:

$$R(t) = at + b \quad (3)$$

where a represents the constant recovery rate and b the initial condition. This model assumes uniform recovery over time and serves primarily as a comparison benchmark.

Exponential Recovery Model:

To capture diminishing recovery rates over time, an exponential model was considered:

$$R(t) = R_{\infty} - Ae^{-kt} \quad (4)$$

where R_{∞} represents the asymptotic recovery level, A the recovery amplitude, and k the recovery constant. This model allows for rapid early recovery followed by slower convergence, a common feature in biological recovery processes.

Power-Law (Scaling) Model:

To explore scale-dependent behavior, a power-law model was evaluated:

$$R(t) = ct^{\alpha} \quad (5)$$

where the exponent α characterizes how recovery scales with time. Power-law behavior is consistent with fractal-like or scale-dependent dynamics, where recovery rates vary depending on the temporal scale of observation.

The applications of these models can be seen in the literature.⁵⁻⁷

Machine Learning Models:

A machine learning model was constructed to evaluate the possibility of forecasting plant recovery dynamics using physiological measurements collected during the recovery phase. In this case, we used a Random Forest regression approach to model nonlinear relationships between time-resolved plant traits and the composite recovery index $R(t)$. We used the characteristics such as plant height, leaf angle, leaf color, turgor, and stress type as features when constructing this model.

When running machine learning models, the dataset was divided into training and testing subsets using a 70/30 split. To improve model robustness and reduce overfitting, 5-fold cross-validation was employed during model training. Random Forest regression performance was evaluated using the coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE).

Statistical Model Comparison:

To compare the above models, we used the R^2 statistics, Residual analysis, and Akaike Information Criterion (AIC) value.

R^2 statistic:

To quantitatively compare candidate recovery models, the coefficient of determination (R^2) was computed to assess the proportion of variance in recovery explained by the model.

$$R^2 = 1 - \frac{\sum_{i=1}^n (R(t_i) - \hat{R}(t_i))^2}{\sum_{i=1}^n (R(t_i) - \bar{R})^2} \quad (6)$$

where $R(t_i)$ denotes the observed recovery metric, $\hat{R}(t_i)$ the model prediction, and $\bar{R}$ the mean recovery value. Higher R^2 values indicate improved explanatory power. However, this metric alone does not capture model adequacy.

Residual Structure:

To evaluate whether model assumptions were satisfied, residuals were analyzed as a function of time:

$$\varepsilon(t_i) = R(t_i) - \hat{R}(t_i) \quad (7)$$

To account for differences in model complexity, models were further compared using the Akaike Information Criterion (AIC).

$$AIC = 2k - 2\ln(\mathcal{L}) \quad (8)$$

Akaike Information Criterion (AIC):

where k is the number of estimated parameters and L is the maximum likelihood of the model given the data. Lower AIC values indicate a preferable balance between goodness-of-fit and model parsimony.

Results and Discussion

Residual plots were inspected for systematic structure, temporal correlation, and heteroscedasticity. Linear models frequently exhibited time-dependent residual patterns, suggesting that recovery dynamics were not well described by a constant-rate process.

In contrast, non-linear models reduced residual structure, indicating improved representation of time-varying recovery behavior.

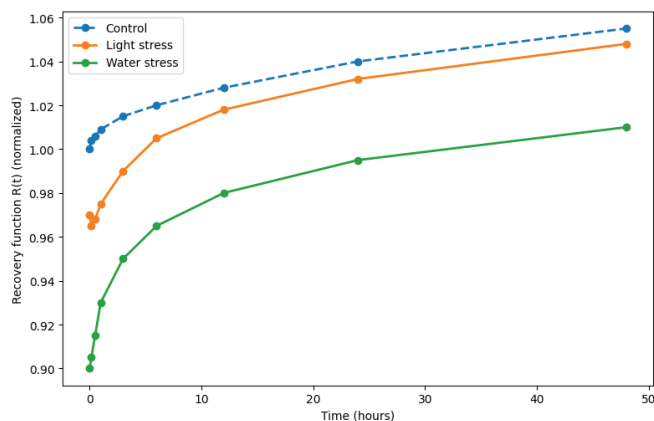


Figure 1: Recovery index of plants over time under different stress conditions. The plot shows the normalized recovery index measured at different time points during the recovery period. The blue line represents the control of plants grown under normal conditions, the orange line represents plants exposed to light stress, and the green line represents plants exposed to water stress. The trajectories illustrate how plant recovery changes over time depending on the type of stress.

As Figure 1 illustrates, control plants show an uninterrupted, steady, and approximately linear trajectory. In contrast, plants that are exposed to stress show a moderate initial decline in recovery immediately after the stress exposure followed by a relatively rapid partial recovery. This pattern suggests that light stress has produced a temporary physiological disruption that

was largely reversible with the resumption of normal conditions.

In contrast, water-stressed plants demonstrated higher deviation of the control group's trajectory, indicating a substantial decline in recovery for both control and light-stressed plants. These results suggest that recovery dynamics are strongly influenced by stress type and are best characterized as time-dependent processes rather than single endpoint outcomes.

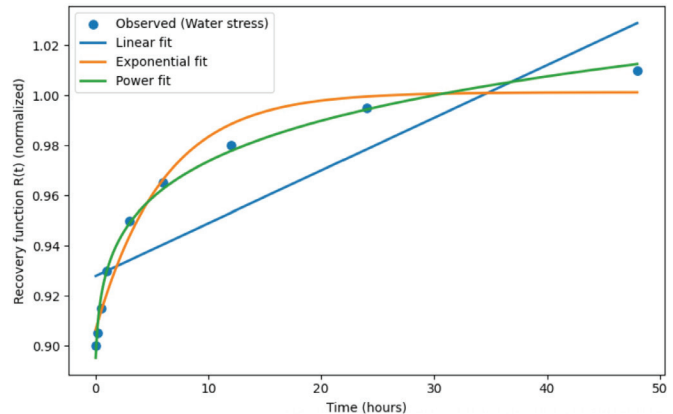


Figure 2: Model fits describing recovery dynamics under water stress. The figure shows the observed recovery index values for water-stressed plants together with the fitted nonlinear models. The plotted curves represent the linear model (blue), exponential model (orange), and power-law model (green), which describe recovery changes over time. The figure allows comparison of how well each model follows the observed recovery trajectory.

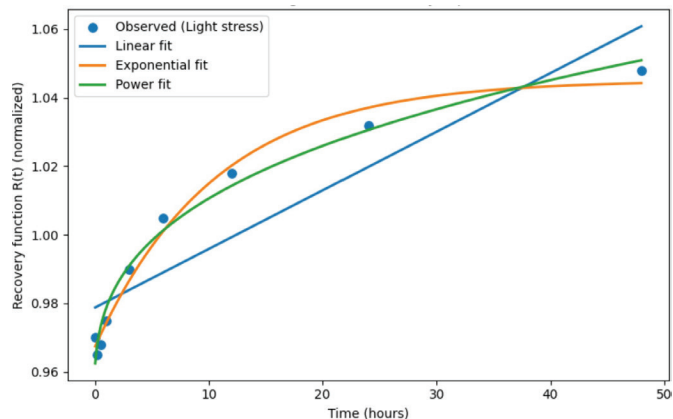


Figure 3: Model fits describing recovery dynamics under light stress. The figure displays the observed recovery index values for plants exposed to light stress along with the fitted nonlinear models. The plotted curves represent the linear (blue), exponential (orange), and power-law (green) models describing the recovery dynamics over time. This visualization allows comparison between the two models in describing the recovery trajectory.

As Figure 2 shows, the observed recovery dynamics of plants under stress clearly follow a nonlinear pattern. The power-law model closely follows the trajectory of the observed recovery dynamics more than the exponential model, indicating that it provides the best fit for the data. Based on Figure 3, the recovery dynamics under light stress are well described by both exponential and power-law models, although the exponential model shows a slight advantage.

As Table 1 shows, the exponential model provides the best fit for the light-stress plants' recovery dynamics, with the lowest AIC value and the highest R^2 value.

In contrast, the power-law model provides the best fit for water-stress recovery. As Table 1 shows, this model records the lowest AIC value and the highest R^2 value, justifying its selection as the preferred model for the recovery trajectory of the water-stress plants.

Furthermore, residual analysis justifies the above model selection. These findings suggest that plants' recovery dynamics after exposure to light and water stress differ fundamentally by stress type, with exponential behavior describing rapid reversible recovery and power-law scaling capturing prolonged recovery following water limitation.

Table 1: The table presents the performance of the linear, exponential, and power-law models for plants exposed to light and water stress. Model performance is evaluated using statistical metrics including the coefficient of determination (R^2), root mean square error (RMSE), and Akaike Information Criterion (AIC), allowing comparison of how well each model fits the recovery data.

Stress	Model	R^2	RMSE	AIC
Light	Linear	0.8129	0.0125	-74.945
Light	Exponential	0.9865	0.0033	-96.629
Light	Power	0.9801	0.0040	-93.096
Water	Linear	0.7120	0.0203	-66.132
Water	Exponential	0.9729	0.0062	-85.414
Water	Power	0.9938	0.0030	-98.664

Based on the forecasting of the plant's recovery index using the Machine Learning model, the results demonstrated that the model successfully captured recovery trends across stress conditions, producing accurate predictions of future recovery trajectories. Figure 4 illustrates the importance of the above features when forecasting the plant's recovery index. Based on the outcomes, stress type, leaf color, leaf angle, and the plant's height are dominant features when forecasting the recovery index. These findings suggest that machine learning methods can support traditional mathematical models by improving the ability to forecast the plant recovery after stress. Such data-driven approaches may help guide agricultural decision-making and support more effective stress management strategies.

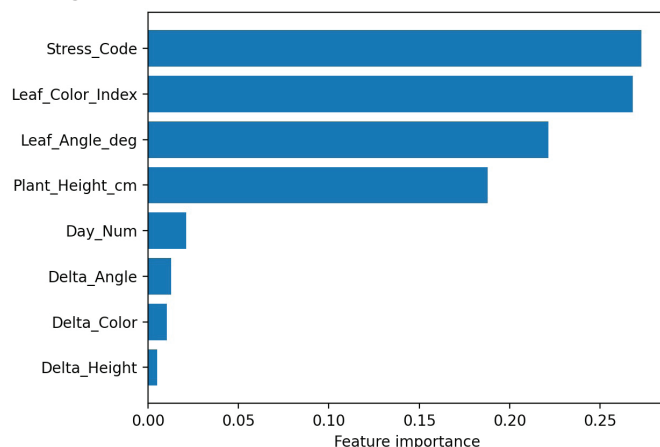


Figure 4: Feature importance for predicting plant recovery index using a Random Forest model. The bar chart shows the relative importance of the variables used in the machine learning model. Each bar represents a predictor variable, including stress type, leaf color, leaf angle, plant height, and turgor. The height of each bar indicates the contribution of that variable to predicting the recovery index.

■ Limitations of the Study

The authors identify several limitations of this study. First, the experimental study was conducted using a single plant species under controlled indoor conditions. This may limit the ability to generalize the results across different plant types or environmental conditions. Secondly, when recording the qualitative physiological measurements, such as leaf color and turgor, observer-based scoring is used despite standardized procedures. This may introduce potential subjectivity. Finally, the observation duration was restricted to a finite recovery window and may not fully describe the long-term adaptation or stress memory effects. Future studies should conduct similar investigations incorporating multiple plant species, extended observation periods, and automated sensing technologies to further improve model robustness and biological generalizability.

■ Conclusion

This study showed that the recovery trajectory of plants experiencing environmental stress, such as light limitation and water deficit, is fundamentally time-dependent and influenced by the type of stress. Based on the experimental findings, it is clear that the linear models are insufficient to capture the recovery dynamics of the plant, where non-linear models provide an improved description of the observed dynamics. Furthermore, exponential behavior best characterizes light-stress recovery, capturing the rapid physiological stabilization after stress removal. For the recovery of water stress, power-law models nicely captured the prolonged and gradual recovery. Furthermore, these results support the interpretation that plants' recovery processes after exposure to environmental stress operate across multiple biological time scales. The integration of machine learning prediction further illustrates the ability to predict plant recovery states using physiological measurements. Future research will extend this framework to multiple plant species to examine whether a generalized recovery model can describe plant resilience across diverse biological systems. This will support practical applications in precision agriculture and environmental stress management.

■ Acknowledgments

The author would like to thank the teachers and mentors whose support and encouragement made this research possible. Their guidance and willingness to share their knowledge played an important role throughout this project. Special thanks are extended to Mr. Rayvan Watson for his mentorship, advice, and continuous support during the research process. His guidance helped strengthen both the direction and quality of this work and was greatly appreciated. The author is also thankful for the supportive academic environment that encouraged curiosity, learning, and independent research.

■ References

1. Lamers, J.; van der Meer, T.; Testerink, C. How Plants Sense and Respond to Stressful Environments. *Plant Physiol.* **2020**, *182* (4), 1624–1635.

2. Osakabe, Y.; Osakabe, K.; Shinozaki, K.; Tran, L.-S. P. Response of Plants to Water Stress. *Front. Plant Sci.* **2014**, *5*, 86.
3. Jacques, C.; Salon, C.; Barnard, R. L.; Vernoud, V.; Prudent, M. Drought Stress Memory at the Plant Cycle Level: A Review. *Plants* **2021**, *10* (9), 1873.
4. Smithers, E. T.; Luo, J.; Dyson, R. J. Mathematical Principles and Models of Plant Growth Mechanics: From Cell Wall Dynamics to Tissue Morphogenesis. *J. Exp. Bot.* **2019**, *70* (14), 3587–3600.
5. Draper, N. R.; Smith, H. *Applied Regression Analysis*; Wiley: New York, 1998.
6. von Bertalanffy, L. Quantitative Laws in Metabolism and Growth. *Q. Rev. Biol.* **1957**, *32*, 217–231.
7. West, G. B.; Brown, J. H.; Enquist, B. J. A General Model for the Origin of Allometric Scaling Laws in Biology. *Science* **1997**, *276*, 122–126.

■ Author

Yenula Wickramasinghe is a sophomore who is curious about how mathematics and machine learning can be used to solve real-world problems. He enjoys exploring data-driven approaches to understanding complex systems and is especially interested in applied mathematics and artificial intelligence.

Orexin-Induced Upregulation of *OX1R* And *BDNF* Suggests an Adaptive Neuroprotective Mechanism Against Oxidative Stress

Youn Joo Lee

Chadwick International, Art Center-daero 97th street 45, Incheon, 22002, South Korea; younjoolee2027@gmail.com

Mentor: Woo Rin Lee

ABSTRACT: Chronic sleep disturbance elevates reactive oxygen species (ROS) and lowers brain-derived neurotrophic factor (BDNF), predisposing the brain to injury. Since orexin (hypocretin) maintains wakefulness and activates stress-responsive pathways, we investigated whether orexin signaling imparts glial resilience to oxidative stress. This study investigated whether lentiviral HCRT overexpression protects A172 glial cells against hydrogen peroxide (H₂O₂)-induced oxidative stress by altering orexin receptor signaling, apoptosis-related responses, and BDNF-associated neuroprotective gene expression. Lentivirus encoding prepro-orexin (*HCRT*) was used to transduce human A172 glial cells. Control groups included empty-virus and no-virus groups. The overexpression of *HCRT* was dose-dependent and was also associated with a significant increase in *OX1R* and *BDNF*. Functionally, orexin enhanced cell viability and diminished apoptosis in response to oxidative stress in a dose-dependent manner, signifying cytoprotection. These synchronized molecular and functional alterations indicate the activation of *OX1R*-dependent pro-survival pathways (cAMP/PKA-CREB, MAPK/ERK, and PI3K-Akt) that augment BDNF-mediated neurotrophic support. The data collectively substantiate an adaptive neuroprotective mechanism in which orexin enhances receptor signaling and trophic tone to mitigate ROS-induced injury. Together, these findings suggest that HCRT may enhance glial resilience by increasing orexin receptor and BDNF-related transcriptional responses, although future validation in additional cell models and protein-level assays will be needed to confirm the broader neuroprotective relevance of this mechanism.

KEYWORDS: Biology, Cell Biology, Orexin (HCRT), A172 Glial Cells.

■ Introduction

Sleep is a biological activity that is important to the health and well-being of humans.¹ Past research has established that sufficient sleep is important to the proper functioning of the immune system.² This implies the existence of a relationship between lack of sleep and a high potential of developing cardiovascular disease, cancer, and depression.³ Besides its impact on the vulnerability to a disease, sleep promotes necessary physiological functions. This beneficial outcome involves hormone control, repairing tissues, and growth. All these functions are critical to physical and mental health. Thus, sleep research and its processes are important for understanding how the body maintains homeostasis.⁴

Orexin, also known as hypocretin, is a protein that may function as a signal sent throughout the brain.⁵ According to a previous study, the neurons in the lateral hypothalamus release orexin A and B, which are neuropeptides—types of chemical messengers within the brain.⁶ Orexin binds to orexin receptors, such as *OX1R*, to promote systems like the noradrenergic system, the histaminergic system, and the cholinergic system, making orexin crucial in sleep-wake regulation in that it promotes wakefulness.⁷ A previous study also supports this by showing that mice lacking orexin exhibited fragmented wakefulness, suggesting that orexin is key for sustaining wakefulness, not just initiating it.⁸

Chronic sleep deprivation triggers cellular stress responses in the human brain. Previous studies indicated that prolonged wakefulness leads to an increase in reactive oxygen species (ROS) in mouse and human brain cells.⁹ This study indicates that a high level of ROS causes damage to cellular components in mitochondria. When the ROS-induced damage is critical, the mitochondria finally release cytochrome C, triggering apoptosis, a cell suicide mechanism.¹⁰ This apoptosis is activated to eliminate the damaged cells in the human brain. However, eventually these consequences lead to neuronal loss and reduced synaptic integrity.¹⁰

Additionally, sleep deprivation has been shown to lower levels of brain-derived neurotrophic factor (*BDNF*).¹¹ This gene encodes a critical protein for neuronal growth, plasticity, and survival. Therefore, previous studies indicate that lower *BDNF* levels may be a key indicator of brain damage in the human brain. Overall, this information suggests that the combined effects of increased ROS, apoptosis, and diminished neurotrophic support.¹² In addition, these cellular alterations may predispose individuals to neurodegenerative processes. Therefore, this mechanism links chronic sleep disturbance to conditions such as Alzheimer's and Parkinson's disease.¹³ Together, the combined effects of increased ROS, apoptosis, and diminished neurotrophic support may create a cellular environment that increases glial vulnerability and contributes to impaired neuronal support during chronic sleep disturbance.

Existing *in vitro* systems of sleep disruption are weak in their capacity to reflect the complex physiology of hyperarousal and disorganized sleep.¹⁴ The available methods are mostly based on cortisol therapy, which is mostly an expression of overall stress reactions. Therefore, this study establishes a novel glia-centered *in vitro* model for examining how HCRT overexpression influences oxidative stress resilience and neuroprotective gene regulation in A172 glial cells. According to this approach, it is possible to simulate hyperarousal-like states more specifically. The approach is also able to identify glia cells, which are important because they are important factors in neuroinflammation, synaptic support, brain homeostasis, etc., but have been neglected in the literature about sleep disturbance. This model proposes an alternative way to study the glial role in neurophysiological mechanisms that are usually discussed as those of neurons by paying attention to orexin-mediated signaling pathways in glia. Finally, this method can be used to assist in offering more specific simulations of hyperarousal-like states, considering the glial cell. To address this gap, the present study develops a glia-centered *in vitro* model of orexin-associated stress signaling by using lentiviral HCRT overexpression in A172 glial cells. This approach differs from conventional cortisol-based stress paradigms by isolating the potential contribution of orexin signaling to oxidative stress resistance, apoptosis regulation, and BDNF-associated neuroprotective responses."

This study is important because it uses lentivirus-mediated orexin overexpression in human glial cells to model sleep disturbance in a new way. This method is different from traditional models that use generalized stress responses because it focuses on molecular markers that are directly related to sleep regulation, like orexin receptor activity and BDNF expression. By looking at downstream effects like oxidative stress and apoptosis, this study provides a clearer and more mechanistic picture of what happens to cells when sleep is disrupted. This research enhances our understanding of glial involvement in sleep physiology and may facilitate the development of orexin-targeted therapies for insomnia and associated disorders.

■ Methods

Cell Culture:

A172 human glioblastoma glial cells (ATCC CRL-1620; A172 cells in Gibco, USA) were grown in Dulbecco Modified Eagle Medium (DMEM, Gibco, USA) with 10% FBS mixed in and 1% penicillin-streptomycin. Cells were kept at 37 °C in a humidified environment with 5 percent CO₂. All the experiments were done on cells between passages 5 and 15.

Lentiviral-Mediated Orexin Overexpression:

A lentiviral virus plasmid with the human prepro-orexin (HCRT) gene under the control of a CMV promoter was acquired commercially (VectorBuilder). This was done through the production of viral particles in HEK293T packaging cells and harvested at 48-72 h after transfecting the cells. The differentiated A172 cells with retinoic acid were transduced with a multiplicity of infection (MOI) of 20 with 80 µg/mL polybrene. Empty lentivirus was given as a control group. The 2

µg/mL of puromycin was used to select the cells after 72 h to maintain steady expression over a period of 5 days. qPCR was used to identify overexpression efficiency.

Quantitative PCR (qPCR):

The A172 cells were lysed using the Total RNA extraction kit (Bioneer). Quantification of RNA was done by Nanodrop, and 1 µg was reverse transcribed using the iScript cDNA synthesis kit (Bio-Rad) and run on a QuantStudio 6 system using SYBR Green Master Mix (Applied Biosystems). Primer sets were specifically made against orexin (HCRT), orexin receptor 1 (OX1R), BDNF, and a reference gene, GAPDH. Relative gene expression was calculated using the 2^{-ΔΔCt} method.

Cell Death and Viability Assay:

The use of an automated cell counter (LUNA-FL™, Logos Biosystems) with dual staining of Acridine Orange/Propidium Iodide (AO/PI) was used to measure cell viability and death. Cells were collected, stained, and analysed instantly to identify living, apoptotic, and dead cells. The outcomes were in the form of the percentage of dead cells relative to the total cell count.

Oxidative stress induction and viability/apoptosis analysis:

To evaluate whether HCRT overexpression altered cellular resilience under oxidative stress, A172 cells were first transduced with lentiviral HCRT or a control vector and allowed to recover for 24 h. Oxidative stress was then induced by treating cells with hydrogen peroxide (H₂O₂) at a final concentration of 100 µM for 24 h. Following H₂O₂ exposure, cell viability and apoptosis were assessed. All groups were processed using the same timing after transduction to ensure direct comparison between control and HCRT-overexpressing cells.

Statistical Analysis:

Every experiment was done in triplicate (n = 3). Results are reported in the form of mean standard deviation (SD). One-way ANOVA and *post hoc* test (Tukey) were used to test statistical significance. A p-value of less than 0.05 was taken as a significant p-value.

■ Results and Discussion

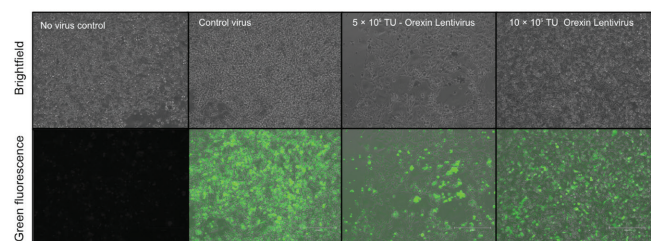


Figure 1: Lentiviral transduction efficiency of Orexin-expressing constructs visualized by GFP fluorescence. Brightfield (top row) and green fluorescence (bottom row) images show A172 cells transduced with increasing doses of Orexin lentivirus. GFP fluorescence increased with viral titer, indicating a clear dose-dependent rise in transduction efficiency. No fluorescence was detected in the no-virus or empty-virus controls, confirming specificity and low background signal. Scale bar = 300 µm.

To evaluate the transduction efficiency of the Orexin-expressing lentivirus, A172 cells were infected with increasing doses of viral particles (5×10^5 and 10×10^5 TU) and examined for GFP fluorescence intensity as a measure of transgene expression. Figure 1 illustrates that both brightfield and green fluorescence images unequivocally demonstrate the successful delivery and expression of lentivirus, exhibiting a dose-dependent efficacy.

The no-virus control and control-virus groups, in contrast, exhibited no detectable fluorescence, thereby confirming the minimal background autofluorescence and the specificity of lentiviral-mediated expression (Figure 1). Cells infected with 5×10^5 TU Orexin lentivirus exhibited moderate GFP signals distributed throughout the culture. In contrast, cells infected with 10×10^5 TU exhibited robust and uniform fluorescence, indicating effective transduction and strong expression of the Orexin gene. Importantly, no cytotoxicity or morphological abnormalities were observed in either infection scenario, indicating the biocompatibility of the lentiviral delivery system.

The rise in GFP fluorescence intensity with higher viral doses shows that the Orexin lentiviral construct works better in human glial-derived A172 cells when the doses are higher. The successful gene delivery achieved with 10×10^5 TU demonstrates that the lentiviral system enabled stable integration and sustained expression without notable cytotoxic effects.

This finding validates a reliable *in vitro* model for subsequent molecular and functional assays investigating Orexin-related neuroprotective signaling. The absence of background signal in the control conditions ensures that GFP expression is exclusively a result of transgene delivery, rather than nonspecific fluorescence or viral artifacts.

These results confirm that lentivirus-mediated overexpression is a viable method for modulating Orexin signaling in neuronal and glial cell models. The established expression gradient (5×10^5 vs. 10×10^5 TU) serves as a basis for subsequent analyses, including qPCR quantification of *HCRT*, *OX1R*, and *BDNF*, to clarify dose-dependent effects on neurotrophic and neuroprotective pathways related to sleep regulation and stress adaptation.

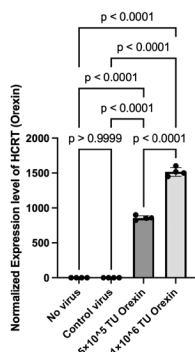


Figure 2: Quantitative PCR analysis of *HCRT* (Orexin) mRNA expression following lentiviral transduction. RT-qPCR measured relative mRNA expression levels of *HCRT* in A172 cells transduced with increasing doses of Orexin lentivirus (5×10^5 and 1×10^6 TU), compared to no-virus and control-virus groups. *HCRT* mRNA was robustly and titratably overexpressed, rising from baseline in controls to $\sim 8.6 \times 10^2$ -fold (5×10^5 TU) and $\sim 1.5 \times 10^3$ -fold (1×10^6 TU). This confirms successful, dose-dependent orexin induction at the transcript level ($n=3$; $p<0.0001$ vs. controls). Data are presented as mean $\pm$ SD ($n=3$).

Lentiviral transduction of A172 cells produced a strong, dose-dependent increase in *HCRT* (Orexin) mRNA (Figure 2). Relative to the control-virus calibrator (set to 1.0), the 5×10^5 TU Orexin group reached $\sim 8.6 \times 10^2$ -fold expression (mean $\pm$ SD: 857 ± 40), and the 1×10^6 TU group reached $\sim 1.5 \times 10^3$ -fold (1517 ± 76), while no-virus and control-virus groups remained at baseline (≈ 0.8 – $1.0 \times$). One-way ANOVA with Tukey's multiple comparisons showed that both Orexin groups were significantly higher than either control ($p < 0.0001$ for all indicated comparisons; $n=3$). These data confirm successful and titratable overexpression of Orexin at the transcript level.

The RT-qPCR results back up the fluorescence results. This data shows that lentiviral delivery results in strong, dose-dependent Orexin expression, whereas controls show no noticeable increase. It is clear that increasing viral input from 5×10^5 to 1×10^6 TU, about twice as much, results in more *HCRT* mRNA. This indicates that expression has not yet reached its maximum in this range. This establishes a reliable model for overexpression that facilitates the investigation of downstream signaling.

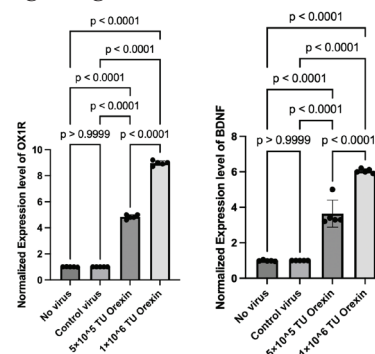


Figure 3: Orexin overexpression upregulates *OX1R* and *BDNF* gene expression in A172 cells. Quantitative RT-qPCR analysis of *OX1R* (left) and *BDNF* (right) mRNA levels following lentiviral transduction of Orexin in A172 cells. Data are presented as mean $\pm$ SD ($n=3$). Orexin overexpression coordinately increased both *OX1R* and *BDNF*, consistent with strengthened receptor signaling and elevated neurotrophic tone. The higher viral dose produced the largest induction (*OX1R* ~ 9 -fold; *BDNF* ~ 6 -fold), supporting a dose-dependent pro-survival molecular response ($n=3$; $p<0.0001$). Statistical analysis was performed using one-way ANOVA with Tukey's multiple comparison test.

Orexin overexpression produced a dose-dependent induction of both *OX1R* and *BDNF* mRNA in A172 cells (Figure 3). Relative to the control-virus calibrator (set to 1.0), the 5×10^5 TU Orexin group showed a ~ 4.8 -fold increase in *OX1R* (mean $\pm$ SD: 4.8 ± 0.2) and a ~ 3.3 -fold increase in *BDNF* (3.3 ± 0.1). The 1×10^6 TU Orexin group further increased expression to ~ 9.0 -fold for *OX1R* (9.0 ± 0.15) and ~ 6.0 -fold for *BDNF* (6.0 ± 0.15). No-virus and control-virus groups remained at baseline. One-way ANOVA with Tukey's multiple comparisons indicated that both Orexin groups were significantly higher than either control for both genes ($p < 0.0001$, all indicated comparisons; $n=3$).

The results show that overexpression of Orexin increases *OX1R* signaling and *BDNF* expression. This supports the idea of an adaptive neuroprotective mechanism in the sleep disturbance model. In the field of sleep-related neurobiology,

elevated Orexin levels are expected to enhance Orexin receptor signaling and activity-dependent trophic pathways. Figure 3 illustrates the simultaneous upregulation of both *OX1R* and *BDNF*. This supports a functional model in which Orexin overexpression can trigger adaptive, neuroprotective mechanisms essential for stress or sleep disturbance conditions. This result shows the numbers needed to link viral dose to molecular readouts. We can then use this system to perform functional tests, such as assessing how well it performs under oxidative stress.

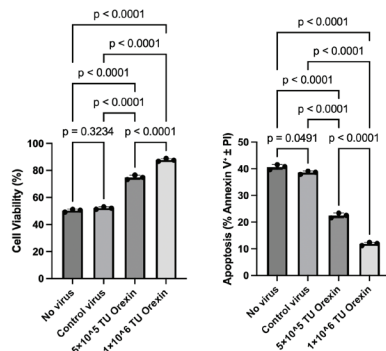


Figure 4: Orexin overexpression confers resistance to oxidative stress. A172 cells transduced with Orexin lentivirus exhibited significantly higher viability and lower apoptosis after H₂O₂ exposure than the no-virus and control-virus groups. Under oxidative stress, orexin significantly improved survival and reduced apoptosis in a dose-dependent manner, demonstrating functional cytoprotection. These effects align with the observed *OX1R*/*BDNF* upregulation, supporting an orexin-linked adaptive neuroprotective axis against ROS-induced injury (n=3; p<0.0001 vs. controls).

Orexin overexpression significantly protected A172 cells from H₂O₂-induced damage (Figure 4). In the viability assay, means (±SD, n=3) rose from ~50–52% in no-virus and control-virus groups to ~75% at 5×10⁵ TU and ~88% at 1×10⁶ TU (one-way ANOVA with Tukey's test, p < 0.0001 for all Orexin vs. control comparisons; no-virus vs. control-virus not significant, p = 0.3234). In the apoptosis assay (Annexin V⁺ ± PI), means declined from ~40–41% (no-virus/control-virus) to ~22% at 5×10⁵ TU and ~12% at 1×10⁶ TU (p < 0.0001 for Orexin vs. controls; a small difference between no-virus and control-virus, p = 0.0491). Together, these data show a clear dose-dependent cytoprotective effect of Orexin.

Given consistent transcriptional increases in *HCRT*, *OX1R*, and *BDNF*, Orexin overexpression may provide neuroprotection against oxidative stress. This may enhance survival and reduce apoptosis in A172 cells. This result indicates that stronger *OX1R* signaling may involve the cAMP/PKA-CREB, MAPK/ERK, and PI3K-Akt pathways, which can activate *BDNF* and other protective programs that prevent cell death. This mechanism protects them from H₂O₂-induced damage.

Although this study did not directly manipulate sleep duration, *HCRT* overexpression was used to model a key molecular feature of sleep disturbance: heightened orexinergic/hypocretin signaling associated with wakefulness, arousal, and stress-related hyperarousal. Orexin neurons promote sustained wakefulness and activate arousal systems; therefore, increasing *HCRT* expression in glial cells provides a reductionist *in vitro* approach for examining how an orexin-enriched cellular envi-

ronment may influence glial resilience under oxidative stress. This model does not reproduce the full systemic physiology of sleep deprivation, but it offers a controlled platform to isolate the relationship between *HCRT* signaling, oxidative stress response, and neurotrophic gene regulation.

Study impact:

This study presents a focused, glia-centric model of sleep disturbance that isolates orexin-driven biology and correlates it with quantifiable cellular protection, enhancing both mechanism and methodology. We demonstrate that lentiviral orexin overexpression enhances *OX1R* and *BDNF* levels while mitigating oxidative-stress injury, thereby establishing a coherent, adaptive neuroprotective axis that is amenable to direct testing. The model serves as a functional platform for evaluating orexin-pathway modulators and *BDNF*/TrkB enhancers, establishing benchmarks for biomarkers of hyperarousal and neuroprotection, and analyzing glial contributions to sleep-related pathology—an aspect frequently neglected in neuron-centric research.¹⁵ Translationally, the results emphasize orexin-targeted interventions as potential treatments for insomnia and stress-related neurodegeneration, and they describe pharmacologic experiments (*OX1R* antagonism and *BDNF* blockade) that can reduce the risk of therapeutic strategies.¹⁶ In general, this work makes things clearer on a mechanistic level and provides a scalable assay framework that speeds up the process of going from cellular insight to therapy design.

Limitations and Future Study:

This work has several limitations that should be addressed in follow-up studies. First, results were derived from an immortalized glial line (A172), which may not recapitulate primary human glia or neuron-glia interactions; future experiments should validate findings in primary astrocytes/microglia, human iPSC-derived co-cultures, brain organoids, and *in vivo* models. Second, the lentiviral overexpression paradigm may exceed physiological orexin levels and introduce MOI/polybrene or integration-site confounds. Therefore, titration to near-physiological expression, CRISPRa/inducible systems, and isogenic controls are warranted. Third, conclusions rely largely on mRNA. In future studies, protein-level and signaling readouts (*OX1R*/*BDNF* protein, CREB-P, ERK, Akt), secreted *BDNF* ELISA, and receptor function assays should be added. Fourth, causal links should be tested using pharmacology and genetics (*OX1R* antagonists such as SB-334867, TrkB blockade with ANA-12, and siRNA/shRNA knock-down of *HCRT*/*OX1R*/*BDNF*). Fourth, the stress paradigm was limited to H₂O₂; expanding to diverse stressors (ER/mitochondrial stress, glutamate excitotoxicity), time-course and single-cell analyses (ROS, ΔΨ_m, caspase-3/7, scRNA-seq/proteomics), and electrophysiology to assess network effects will clarify the breadth, mechanism, and translational relevance of the proposed neuroprotective axis. Finally, the present study evaluates *OX1R* and *BDNF* primarily at the mRNA level using RT-qPCR. Because transcript abundance does not always predict protein abundance, the results should be interpreted as

evidence of transcriptional regulation rather than direct confirmation of receptor or neurotrophic protein changes. Future studies should include Western blotting or ELISA to determine whether HCRT-induced increases in OX1R and BDNF mRNA are accompanied by corresponding protein-level changes and to strengthen the proposed neuroprotective mechanism.

■ Conclusion

The study was successfully able to establish a glial-cell model using the lentivirus-mediated Orexin (HCRT) overexpression to study its reaction to stress conditions. Resultantly, this model showed increased Orexin, OX1R, and BDNF expression. In addition, it was observed that cell viability increased while apoptosis decreased under oxidative stress, showing increased cell survival and resilience.

Such molecular and functional outcomes suggest Orexin activates neuroprotective signaling pathways for increased resilience under stress. Specifically, Orexin likely engages OX1R-dependent cascades, such as cAMP/PKA-CREB, MAPK/ERK, and PI3K-Akt, all of which are pathways that support the plasticity of cells. In our observation, activation of these pathways led to BDNF-mediated neuronal survival and stress resistance.

Such mechanisms suggest that orexin signaling represents an adaptive response to sleep disturbance or oxidative injury. By increasing cell viability, reducing apoptosis, and expression of BDNF, a crucial factor for survival, Orexin appears to have a role in maintaining integrity under stress conditions such as sleep deprivation or hyperarousal. In conclusion, lentivirus-mediated Orexin overexpression offers a reliable platform to study cellular mechanisms underlying sleep disturbance. It also underscores the potential for Orexin-targeted pathways to be used as therapeutic avenues for insomnia and neurodegenerative conditions due to Orexin's role in enhancing neuronal survival and stress resilience.

■ Acknowledgments

I would like to thank Dr. Lee for his support and guidance for this research.

■ References

- Ramar, K.; Malhotra, R. K.; Carden, K. A.; Martin, J. L.; Abbasi-Feinberg, F.; Aurora, R. N.; Kapur, V. K.; Olson, E. J.; Rosen, C. L.; Rowley, J. A.; Shelgikar, A. V.; Trotti, L. M. Sleep Is Essential to Health: An American Academy of Sleep Medicine Position Statement. *J. Clin. Sleep Med.* **2021**, *17* (10), 2115–2119. DOI: 10.5664/jcsm.9476.
- Besedovsky, L.; Lange, T.; Born, J. Sleep and Immune Function. *Pflügers Arch.* **2012**, *463* (1), 121–137. DOI: 10.1007/s00424-011-1044-0.
- Javaheri, S.; Redline, S. Insomnia and Risk of Cardiovascular Disease. *Chest* **2017**, *152* (2), 435–444. DOI: 10.1016/j.chest.2017.01.026.
- Deboer, T. Sleep Homeostasis and the Circadian Clock: Do the Circadian Pacemaker and the Sleep Homeostat Influence Each Other's Functioning? *Neurobiol. Sleep Circadian Rhythms* **2018**, *5*, 68–77. DOI: 10.1016/j.nbscr.2018.02.003.
- Inutsuka, A.; Yamanaka, A. The Physiological Role of Orexin/Hypocretin Neurons in the Regulation of Sleep/Wakefulness and Neuroendocrine Functions. *Front. Endocrinol. (Lausanne)* **2013**, *4*, 18. DOI: 10.3389/fendo.2013.00018.
- Ferrari, L. L.; Park, D.; Zhu, L.; Palmer, M. R.; Broadhurst, R. Y.; Arrigoni, E. Regulation of Lateral Hypothalamic Orexin Activity by Local GABAergic Neurons. *J. Neurosci.* **2018**, *38* (6), 1588–1599. DOI: 10.1523/JNEUROSCI.1925-17.2017.
- Scammell, T. E.; Winrow, C. J. Orexin Receptors: Pharmacology and Therapeutic Opportunities. *Annu. Rev. Pharmacol. Toxicol.* **2011**, *51*, 243–266. DOI: 10.1146/annurev-pharmtox-010510-100528.
- Diniz Behn, C. G.; Klerman, E. B.; Mochizuki, T.; Lin, S.-C.; Scammell, T. E. Abnormal Sleep/Wake Dynamics in Orexin Knockout Mice. *Sleep* **2010**, *33* (3), 297–306. DOI: 10.1093/sleep/33.3.297.
- Davinelli, S.; Medoro, A.; Savino, R.; Scapagnini, G. Sleep and Oxidative Stress: Current Perspectives on the Role of NRF2. *Cell. Mol. Neurobiol.* **2024**, *44* (1), 52. DOI: 10.1007/s10571-024-01487-0.
- Liu, N.; Liu, Y.; Wang, Y.; Feng, C.; Piao, M.; Liu, M. Oxidative Cell Death in the Central Nervous System: Mechanisms and Therapeutic Strategies. *Front. Cell Dev. Biol.* **2025**, *13*, 1562344. DOI: 10.3389/fcell.2025.1562344.
- Monteiro, B. C.; Monteiro, S.; Candida, M.; Adler, N.; Paes, F.; Rocha, N.; Nardi, A. E.; Murillo-Rodriguez, E.; Machado, S. Relationship Between Brain-Derived Neurotrophic Factor (Bdnf) and Sleep on Depression: A Critical Review. *Clin. Pract. Epidemiol. Ment. Health* **2017**, *13*, 213–219. DOI: 10.2174/1745017901713010213.
- Miranda, M.; Morici, J. F.; Zannoni, M. B.; Bekinschtein, P. Brain-Derived Neurotrophic Factor: A Key Molecule for Memory in the Healthy and the Pathological Brain. *Front. Cell. Neurosci.* **2019**, *13*, 363. DOI: 10.3389/fncel.2019.00363.
- Owen, J. E.; Veasey, S. C. Impact of Sleep Disturbances on Neurodegeneration: Insight from Studies in Animal Models. *Neurobiol. Dis.* **2020**, *139*, 104820. DOI: 10.1016/j.nbd.2020.104820.
- Kalmbach, D. A.; Anderson, J. R.; Drake, C. L. The Impact of Stress on Sleep: Pathogenic Sleep Reactivity as a Vulnerability to Insomnia and Circadian Disorders. *J. Sleep Res.* **2018**, *27* (6), e12710. DOI: 10.1111/jsr.12710.
- Schirò, G.; Iacono, S.; Ragonese, P.; Aridon, P.; Salemi, G.; Balistreri, C. R. A Brief Overview on BDNF-Trk Pathway in the Nervous System: A Potential Biomarker or Possible Target in Treatment of Multiple Sclerosis? *Front. Neurol.* **2022**, *13*, 917527. DOI: 10.3389/fneur.2022.917527.
- Żelabowski, K.; Petrov, W.; Wojtysiak, K.; Ratka, Z.; Biedka, K.; Wesołowski, M.; Fus, K.; Ślebioda, D.; Rusinek, M.; Sterkowicz, M.; Radzka, I.; Chłopaś-Konowalek, A. Targeting the Orexin System in the Pharmacological Management of Insomnia and Other Diseases: Suvorexant, Lemborexant, Daridorexant, and Novel Experimental Agents. *Int. J. Mol. Sci.* **2025**, *26* (17). DOI: 10.3390/ijms26178700.

■ Author

Youn Joo Lee is an 11th-grade student at the Chadwick International School in Incheon, South Korea. She is interested in engineering and neuroscience, especially regarding sleep medicine, and intends to pursue a related field at university.

A Pilot Investigation of Antibiotic Susceptibility Profile of *Escherichia coli* in Wastewater from a Boarding High School in Türkiye

Zeynep M. Çelik

Robert College, Kuruçeşme Cad.No:87, Arnavutköy, Beşiktaş, İstanbul, 34345, Türkiye; zeynepmirayc08@gmail.com
Mentor: Dünya Önen

ABSTRACT: This pilot study investigates the microbiological quality and antibiotic resistance profiles of *Escherichia coli* isolates from domestic wastewater at a boarding high school where the population is 1500 people in İstanbul, Türkiye. Samples collected from the main sewer line were subjected to enrichment, isolation on TBX (Tryptone Bile X-Glucuronide) agar, and Kirby-Bauer disk diffusion tests. The sensitivity of the isolates to Ampicillin, Tigecycline, Ciprofloxacin, and Cephalothin was evaluated according to Clinical and Laboratory Standards Institute (CLSI)'s 2024 standards. In this preliminary sample, the two presumptive *E. coli* isolates tested were susceptible to the antibiotics evaluated. These results indicate that the burden of resistant bacteria in healthy and educational populations is lower than in hospital-derived or general municipal wastewaters, reflecting manageable local resistance dynamics.

KEYWORDS: Microbiology, Antibiotic Resistance, *Escherichia coli*, Wastewater, Public Health.

■ Introduction

Antimicrobial resistance is a silent pandemic threatening the protectors of modern medicine and is considered one of the most critical global health challenges of the 21st century.¹ Bacterial resistance mechanisms to antibiotics lead to untreatable infections, increased healthcare costs, and high mortality rates. Data published by the World Health Organization (WHO) underscores the importance of monitoring resistance trends not only in clinical samples but also in environmental reservoirs.²

Wastewater is identified as a "hotspot" for the spread of antibiotics and the transfer of resistance genes between different bacterial species through horizontal gene transfer.^{3,4} Wastewater combines domestic, industrial, and sometimes hospital waste, creating an ideal selective pressure environment for resistant bacteria.⁵ *Escherichia coli* is a globally accepted "indicator organism" for monitoring antimicrobial resistance because it is a natural member of the mammalian gut flora and a primary indicator of fecal contamination.^{6,7}

Antibiotic discs from β -lactam, tetracycline, and fluoroquinolone groups, consisting of Ampicillin, Cephalothin, Tigecycline, and Ciprofloxacin, respectively, were used because they represent different mechanisms of action and antibiotic classes. The increasing rates of bacterial antimicrobial resistance worldwide necessitate the evaluation of susceptibility profiles to different antibiotic groups. According to World Health Organization (WHO) data, resistance rates to β -lactams and fluoroquinolones can exceed 20–50% in many clinical pathogens. In contrast, tigecycline stands out as an alternative agent that largely maintains its efficacy against multidrug-resistant strains. Therefore, the combined study of these antibiotics, representing the β -lactam (ampicillin and cephalothin), fluoro-

quinolone (ciprofloxacin), and tetracycline (tigecycline) groups, allows for a comprehensive evaluation of the susceptibility and resistance characteristics of *E. coli* to antibiotics with different mechanisms of action.²²

This pilot study aims to contribute to local public health data by determining the resistance load of wastewater in our school campus. The antibiotics tested have different mechanisms of action and are agents frequently used in clinical practice.

■ Methods

On February 18th, at 9:30 AM, with prior permission from the school administration, a 500 mL composite sample (2 bottles, 250 mL) was taken from the main sewer outlet of the entire school and boarding sections; sterile bottles were simultaneously used to collect 500 mL of sample from the sewage using the immersion method at one time. These samples were placed in a portbag containing ice packs at +4°C in the wake of collection of sample. The samples were transported to the laboratory immediately under cold chain conditions, as seen in Figures 1 and 2.



Figure 1: The main sewage outlet from where the sample was taken. This sampling point ensures that the collected samples reflect the overall resistance profile of the campus because it represents the combined wastewater of the entire school population.

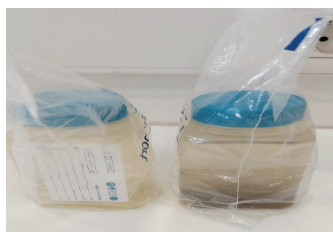


Figure 2: Containers holding samples of wastewater from a high school. Proper collection and transport under controlled conditions preserved sample integrity.

E. coli isolation was performed according to ISO 7251:2005 (ISO, 2005).²⁰ 5 mL of each sample was homogenized in 225 mL of sterile peptone water. From the homogenized homogenates, 1 mL subsample was taken and added to 9 mL of Tryptone Soy Broth (TSB) containing 20 mg novobiocin per liter, and pre-enriched at 37°C for 24 hours. On Day 2, Thermotolerant coliform selection was then performed using Lauryl Sulphate Tryptone (LST) broth and Brilliant Green Bile Broth (BGBB). Incubation at 44°C for 24 hours during the enrichment and isolation stages was preferred for the selective development of thermotolerant (fecal) coliforms. On Day 3, *E. coli* forms typical colonies at 44°C due to gas formation in the BGBB, as named presumptive *E. coli*. 100µL samples were taken from the cultures and inoculated onto Tryptone Bile X-glucuronide Agar (TBX) using the spread plate method, and then incubated for 48 hours at 44°C. After incubation on Day 4, colonies showing a metallic green reflection on TBX were used to confirm β -glucuronidase activity on presumptive *E. coli* isolates. This temperature suppresses the growth of other mesophilic bacteria that may be present in the medium, facilitating the isolation of the target microorganism, as demonstrated in Figures 3 and 4.⁹ Samples were inoculated onto TBX chromogenic agar and incubated at 44°C for 48 hours. Characteristic blue-green *E. coli* colonies were purified, as seen in Figure 5.^{3,10}

Two suspected *E. coli* were isolated before being passed through TBX the night before, were collected using a loop, and placed into tubes containing Tryptone Broth with 1% tryptophan for the indole test, and into tubes containing methyl red Voges-Proskauer (MR-VP) medium (double tubes) for the methyl red and Voges-Proskauer tests. After incubating the tubes overnight, the formation of a bright red/purple ring at the top of the tube for the indole test was considered positive. For the methyl red test, methyl red indicator was added to the MR-VP tube, and colonies showing a red color in the tubes were considered *E. coli* positive. Barrit's reagent A (0.6 ml of 5% α -naphthol solution in alcohol) and Barrit's reagent B (0.2 ml of 40% KOH solution and 2 drops of 0.5% creatine solution) were added to the other MR-VP tubes, and those that turned red were considered *E. coli* positive.¹⁹



Figure 3: Selection of blue blue-green colonies in TBX agar before using an inoculation loop. The appearance of blue blue-green colonies confirms the successful isolation of *E. coli*.

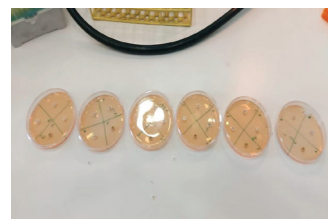


Figure 4: Muller Hinton Agar plates before incubation.

In the wake of biochemical confirmation of *E. coli*, two colonies were taken from the culture and suspended in physiological saline according to the McFarland 0.5 (1×10^8 cfu/ml) opacity standard. Then, 0.1 mL smear plates were made on Mueller-Hinton agar. The plated Petri dishes were kept in a cabinet at room temperature for 5-10 minutes before incubation. After this stage, 4 antibiotic discs, ampicillin (10 µg), ciprofloxacin (5 µg), Tigecycline (15 µg), and cephalothin (30 µg), were placed in one Petri dish. After the inoculation process was completed, 6 Petri dishes were incubated at 37°C for 24 hours, and antibiotic resistance was evaluated according to Clinical and Laboratory Standards Institute (CLSI) 2024 standards by measuring the zone diameters as seen in Table 1. At the end of this period, inhibition zones were formed depending on the susceptibility of the enriched *E. coli* colonies to the four antibiotics. The formation and measurement of inhibition zones are illustrated in Figure 5. Ampicillin, Tigecycline, Ciprofloxacin, and Cephalothin resulted in inhibition zones with average diameters of 26 mm, 20 mm, 28 mm, and 18 mm, respectively, as seen in Table 1. According to the CLSI 2024 Antimicrobial Susceptibility Testing Guidelines, the grown *E. coli* was observed to be susceptible to all four antibiotics. This research comprehensively analyzes the antibiotic susceptibility profiles of *Escherichia coli* strains isolated from wastewater on a boarding school campus, considering global resistance trends, biochemical mechanisms, population dynamics, seasonal factors, and methodological limitations. Our key finding, "almost complete susceptibility," may refute the initial hypothesis that "resistant bacteria will be detected," but provides highly unique and strategic data when compared to existing literature on the microbiological health of school ecosystems.

For quality control purposes in antibiotic susceptibility testing, the *Escherichia coli* ATCC 25922 reference strain was used. The control strain was tested under the same conditions (0.5 McFarland suspension, same medium, and incubation

conditions) on each test day to verify that the inhibition zone diameters were within the range specified in the CLSI 2024 standards. The test results showed that the zone diameters of the control strain were within acceptable limits; thus, the reliability of the findings obtained for the isolates was confirmed. The Kirby-Bauer disk diffusion method was applied. Inhibition zones were measured and interpreted according to the CLSI 2024 standards.²¹

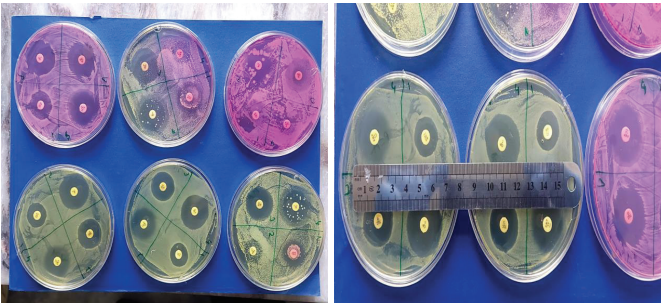


Figure 5: A ruler was placed on two Muller-Hinton Agar plates after 24 hours of incubation at 37°C to measure the diameters of the inhibition zones. Clearly visible inhibition zones show antibiotic activity against the isolates, supporting the finding of full susceptibility across every antibiotic that has been tested.

■ Results and Discussion

Results:

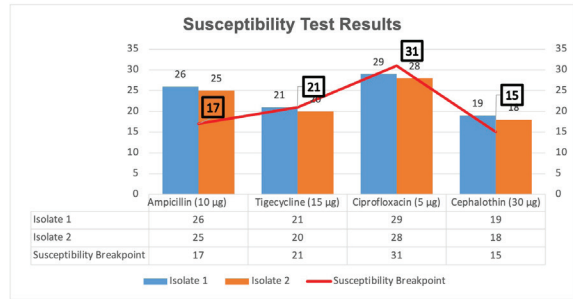
Samples were collected on February 18th at 9:30 AM and immediately taken to the laboratory for analysis. After the *E. coli* identification analysis, at the end of the sixth day, two presumptive strains were confirmed as *E. coli* strains and were transferred to Muller-Hinton Agar, and four antibiotic discs were placed to observe the antibiotic reactions against the *E. coli* strains; the resulting diameters were measured with a ruler.

The data obtained from the analyses showed that the isolated *E. coli* strains were susceptible and intermediately susceptible to all four different classes of antibiotics tested.

Table 1: Inhibition zone diameters of isolates (mm). All measured inhibition zones exceed the CLSI susceptibility breakpoints (mm).

Isolates #	Antibiotics															
	Interpretive Categories and Zone Diameter Breakpoints, nearest whole mm.(CLSI,2024)															
	Ampicillin (10 µg)				Tigecycline (15 µg)				Ciprofloxacin (5 µg)				Cephalothin (30 µg)			
	S	SSD	I	R	S	SSD	I	R	S	SSD	I	R	S	SSD	I	R
#1 (mm)	26 (S)				21 (S)				29 (I)				19 (S)			
#2 (mm)	25 (S)				20 (I)				28 (I)				18 (S)			

S (Susceptible), SSD (Susceptible-Dose Dependent), I (Intermediate), R (Resistant).



Discussion:

Our susceptibility rate contrasts sharply with a vast body of literature arguing that urban wastewater is a primary reservoir for resistant pathogens. To understand the reasons for this, comparison with studies in geographies where resistance is much higher is essential. Research in densely populated urban areas like Dhaka, Bangladesh, where antibiotic use is uncontrolled, shows that wastewater acts like a "resistance factory." A study by Siddique *et al.* found that over 90% of *E. coli* strains isolated from urban wastewater showed resistance to at least three different classes of antibiotics (MultiDrug Resistance MDR).⁵ In the case of Dhaka, resistance to ampicillin and ciprofloxacin, in particular, has reached endemic levels. The detection of 0% resistance to these antibiotics in our study data demonstrates the extent to which the settlement is protected from urban pollution pressures and the introduction of resistant strains. Our results are remarkable even when compared to data from developed countries. Studies conducted in 17 municipal wastewater treatment plants in Oregon, USA, found ampicillin resistance to be in the 30-40% range and ciprofloxacin resistance in the 10-15% range.³ The absence of this baseline resistance level, seen even in regions with well-organized healthcare systems like Oregon, on our high school campus can be attributed to both the limited size of the population and the high level of awareness among campus residents regarding antibiotic use. The prevalence of resistant bacteria is directly related to the volume and quality of antibiotic consumption in that community.² The 0% resistance rate in our study indicates that the gut flora (microbiome) of individuals within the campus is dominated by "wild-type" susceptible bacteria.¹¹ Furthermore, urban wastewater systems often combine hospital, industrial, and domestic waste. The isolation of our school's sewage system from high-resistance foci such as the hospital eliminates the risk of "cross-contamination."¹⁵ This isolated micro-ecosystem lacks the biochemical pressure necessary for the establishment and spread of resistant strains. The fact that the study was conducted on February 18th, during the winter period, plays a central role in interpreting the results. Temperature is a variable that directly affects bacterial metabolic rate, reproduction rate, and, most importantly, horizontal gene transfer (conjugation).¹² In a study conducted on municipal wastewater in Poland, it was found that resistance rates in *E. coli* strains isolated during the winter months were significantly lower than those isolated during the summer months.¹³ Low temperatures suppress the transfer of plasmids carrying resistance genes between bacteria (conjugation frequency). The low temperatures in February may have minimized the rate of bacterial genetic exchange in the sewer line. In addition, the flow rate of water in the sewer system increases with precipitation during the winter months, which dilutes the bacterial concentration (dilution effect).³ This dilution could technically reduce the likelihood of detecting resistant bacteria. If this pilot study were repeated during the summer months, a low level of resistance might be detected due to accelerated horizontal gene transfer resulting from increased temperatures and decreased water flow. Therefore, the

detected full susceptibility result may reflect the period when the population is seasonally at its "cleanest".

Modern public health understanding is based on the "One Health" approach, which argues that human health and environmental health are inseparable.¹⁴ When wastewater is discharged into natural water sources without treatment, it carries resistance elements into the environment.^{4,12} The fact that our high school's wastewater would be clean in terms of resistance means that resistant pathogens are not being pumped into the local ecosystem. Environmental bacteria (aquatic or soil bacteria) can become "genetic reservoirs" by incorporating resistance genes from wastewater.¹⁵ Our study may prove that the school campus does not negatively contribute to this global resistance chain. This may be an indicator of the success of educational institutions in terms of environmental sustainability.¹⁶

Wastewater-based surveillance is like a massive biobank where the biological data of thousands of individuals are collected in a single sample.¹⁷ A sample taken from the school's sewage outlet provided a summary of the common "gut health" of all individuals living on that campus. This approach is the most effective way to monitor a community's resistance profile when individual testing is costly. A study conducted in Sweden by Hutinel *et al.* demonstrated that wastewater analyses showed more than 80% agreement with clinical resistance data in that area.⁶ The exact sensitivity in our high school would likely parallel infection data and antibiotic prescribing rates in the school's infirmary records. This provides school administrations with a scientific basis for developing public health policies. The results of this pilot study can be considered a concrete reflection of the health education and awareness of rational drug use provided at the school. The absence of resistant bacteria is also indirect evidence that hygiene standards and food safety (cafeteria, etc.) within the school are high. This is because resistant bacteria usually infiltrate the community through contaminated food or inadequate sanitation.¹⁰

■ Conclusion

In conclusion, this pilot study, conducted on February 18, 2026, demonstrated that wastewater on our high school campus currently exhibits a low-risk profile in terms of antibiotic resistance. The susceptibility of the isolated *Escherichia coli* strains to Ampicillin, Tigecycline, Ciprofloxacin, and Cephalothin antibiotics was confirmed by the inhibition zones being well above the CLSI 2024 standards. The rejection of the initial hypothesis that "resistant bacteria will be detected" reveals that the school community may have a healthy microbial flora and that the campus ecosystem may not yet be severely impacted by the global antimicrobial resistance crisis. However, given the dynamic nature of antimicrobial resistance, maintaining this positive picture requires the continued commitment to periodic surveillance studies, rational antibiotic use policies, and "One Health" focused environmental monitoring activities.^{14,18}

This research is a successful pilot study aimed at raising environmental and public health awareness in our school. Furthermore, it is believed that future studies in this area would

benefit from collecting samples at different times and from different depths to detect seasonal fluctuations and conducting genetic research to identify potential resistant microorganism genes.

■ Acknowledgments

I would like to express my sincere gratitude to my project advisor and AP Biology teacher, Dünya Önen. Without her continuous support, guidance, and invaluable feedback, this research would not have been possible. I would also like to thank the school authorities for their support and for facilitating my access to the wastewater samples from the sewage system.

■ References

- Centers for Disease Control and Prevention, CDC (2024). *About Antimicrobial Resistance*. Antimicrobial Resistance. <https://www.cdc.gov/antimicrobial-resistance/about/index.html>
- Ajulo, S., & Awosile, B. (2024). Global antimicrobial resistance and use surveillance system (GLASS 2022): Investigating the relationship between antimicrobial resistance and antimicrobial consumption data across the participating countries. *PLoS ONE*, 19(2). <https://doi.org/10.1371/journal.pone.0297921>
- Khorshidi-Zadeh, M., Yiu, S. Y., Nguyen, J. N., Garza, G. L., Waite-Cusic, J., Radniecki, T. S., & Navab-Daneshmand, T. (2021). Antibiotic resistance profile of isolates in 17 municipal wastewater utilities across Oregon. *medRxiv (Cold Spring Harbor Laboratory)*. <https://doi.org/10.1101/2021.11.15.21266366>
- Mukherjee, M., Laird, E. D., Gentry, T. J., Brooks, J. P., & Karthikeyan, R. (2021). Increased Antimicrobial and Multidrug Resistance Downstream of Wastewater Treatment Plants in an Urban Watershed. *Frontiers in Microbiology*, 12. <https://doi.org/10.3389/fmicb.2021.657353>
- Siddique, A. B., Munni, A., Hasan, M., Raj, R., Mutalib, Md. A., Sikder, Md. T., Okino, T., Ahmed, A., & Hossain, M. S. (2024). Investigation and detection of multiple antibiotic-resistant pathogenic bacteria in municipal wastewater of Dhaka city. *Discover Water*, 4(1). <https://doi.org/10.1007/s43832-024-00114-9>
- Hutinel, M., Huijbers, P. M. C., Fick, J., Åhrén, C., Larsson, D. G. J., & Flach, C. (2019). Population-level surveillance of antibiotic resistance in *Escherichia coli* through sewage analysis. *Eurosurveillance*, 24(37).
- Thani, T. S., Symekher, S., Boga, H., & Oundo, J. (2016). Isolation and characterization of *Escherichia coli* pathotypes and factors associated with well and borehole water contamination in Mombasa County. *Pan African Medical Journal*, 23, 12. <https://doi.org/10.11604/pamj.2016.23.12.7755>
- Suzuki, Y., Hashimoto, R., Xie, H., Nishimura, E., Nishiyama, M., Nukazawa, K., & Ishii, S. (2019). Growth and antibiotic resistance acquisition of *Escherichia coli* in a river that receives treated sewage effluent. *The Science of The Total Environment*, 690, 696. <https://doi.org/10.1016/j.scitotenv.2019.07.050>
- ISO 16649-2 (2001). International Standard, Microbiology of food and animal feeding stuffs-Horizontal method for the Enumeration of β -glucuronidase-positive *Escherichia coli*.
- Summerlin, H. N., Pola, C. C., McLamore, E. S., Gentry, T., Karthikeyan, R., Gomes, C. L., Summerlin, H. N., Pola, C. C., McLamore, E. S., Gentry, T., Karthikeyan, R., & Gomes, C. L. (2021). Prevalence of *Escherichia coli* and Antibiotic-Resistant Bacteria During Fresh Produce Production (Romaine Lettuce) Using Municipal Wastewater Effluents. *Frontiers in Microbiology*, 12, 660047. <https://doi.org/10.3389/fmicb.2021.660047>

11. Sudhakari, P. A., & Ramisetty, B. C. M. (2022). Resistome prevalence and diversity in *Escherichia coli* isolates of global wastewaters. *Research Square (Research Square)*. <https://doi.org/10.21203/rs.3.rs-2048981/v1>
12. Larsson, D. G. J., & Flach, C. (2021). Antibiotic resistance in the environment. *Nature Reviews Microbiology*, 20(5), 257. <https://doi.org/10.1038/s41579-021-00649-x>
13. Osińska, A., Korzeniewska, E., Harnisz, M., Niestępski, S., & Jachimowicz, P. (2019). The occurrence of antibiotic-resistant bacteria, including *Escherichia coli*, in municipal wastewater and river water. *E3S Web of Conferences*, 100, 61. <https://doi.org/10.1051/e3sconf/201910000061>
14. Franklin, A. M., Weller, D. L., Durso, L. M., Bagley, M., Davis, B. C., Frye, J. G., Grim, C. J., Ibekwe, A. M., Jahne, M. A., Keely, S. P., Kraft, A. L., McConn, B. R., Mitchell, R. M., Ottesen, A., Sharma, M., Strain, E., Tadesse, D. A., Tate, H., Wells, J. E., ... Garland, J. L. (2024). A one health approach for monitoring antimicrobial resistance: developing a national freshwater pilot effort. *Frontiers in Water*, 6. <https://doi.org/10.3389/frwa.2024.1359109>
15. Nguyễn, Q. A., Vu, H. P., Nguyen, L. N., Wang, Q., Djordjevic, S. P., Donner, E., Yin, H., & Nghiem, L. D. (2021). Monitoring antibiotic resistance genes in wastewater treatment: Current strategies and future challenges [Review of *Monitoring antibiotic resistance genes in wastewater treatment: Current strategies and future challenges*]. *The Science of The Total Environment*, 783, 146964. Elsevier BV. <https://doi.org/10.1016/j.scitotenv.2021.146964>
16. Dávidová-Geržová, L., Lausova, J., Sukkar, I., Nešporová, K., Nechutna, L., Vlková, K., Chudějová, K., Krütová, M., Palkovičová, J., Kaspar, J., & Dolejská, M. (2023). Hospital and community wastewater as a source of multidrug-resistant ESBL-producing *Escherichia coli*. *Frontiers in Cellular and Infection Microbiology*, 13. <https://doi.org/10.3389/fcimb.2023.1184081>
17. Chau, K., Barker, L., Budgell, E. P., Vihta, K. D., Sims, N., Kasprzyk-Hordern, B., Harriss, E., Crook, D. W., Read, D. S., Walker, A. S., & Stoesser, N. (2022). Systematic review of wastewater surveillance of antimicrobial resistance in human populations. *Environment International*, 162, 107171. <https://doi.org/10.1016/j.envint.2022.107171>
18. Tiwari, A., Kurittu, P., Al-Mustapha, A. I., Heljanko, V., Johansson, V., Thakali, O., Mishra, S. K., Lehto, K., Lipponen, A., Oikarinen, S., Pitkänen, T., & Heikinheimo, A. (2022). Wastewater surveillance of antibiotic-resistant bacterial pathogens: A systematic review [Review of *Wastewater surveillance of antibiotic-resistant bacterial pathogens: A systematic review*]. *Frontiers in Microbiology*, 13. Frontiers Media. <https://doi.org/10.3389/fmicb.2022.977106>
19. Roberts D, Greenwood M. (2003). Practical Food Microbiology. 3rd edition, Blackwell Publishing, 150-170.
20. ISO 7251:2005, Microbiology of food and animal feeding stuffs — Horizontal method for the detection and enumeration of presumptive *Escherichia coli* — Most probable number technique. Edition 3, (2005). <https://www.iso.org/standard/34568.html>.
21. Clinical and Laboratory Standards Institute, CLSI (2024). <https://clsi.org/standards/m100>.
22. World Health Organization, WHO (2025). Global antibiotic resistance surveillance report 2025.

■ Author

Zeynep Miray Çelik is a junior at American Robert College of Istanbul. She is highly interested in pursuing molecular and cellular biology or microbiology as a major in college. She hopes to pursue medicine as a career after earning her bachelor's degree.



Marine Biology Research at Bahamas

Unique and exclusive partnership with the Gerace Research Center (GRC) in San Salvador, Bahamas to offer marine biology research opportunities for high school teachers and students.

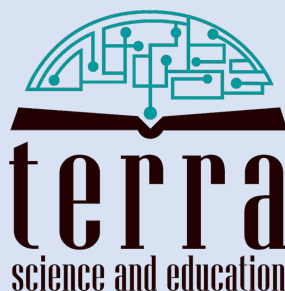
- Terra has exclusive rights to offer the program to high school teachers and students around world.
- All trips entail extensive snorkeling in Bahamian reefs as well as other scientific and cultural activities.
- Terra will schedule the program with GRC and book the flights from US to the GRC site.
- Fees include travel within the US to Island, lodging, meals, and hotels for transfers, and courses.
- For more information, please visit terraed.org/bahamas.html

Terra is a N.Y. based 501.c.3 non-profit organization dedicated for improving K-16 education

IJHSR

International
Journal of
High School
Research

is a publication of



N.Y. based 501.c.3 non-profit organization
dedicated for improving K-16 education