Pakistan Journal of Statistics and Operation Research

A Generalized Exponential Regression Model for Predicting High-Grade Glioma Growth in Paediatric Patients

Md Shohel Rana[1,3*], Mohana Sundaram Muthuvalu[2]

* Corresponding Author

1. Department of Applied Science, Universiti Teknologi PETRONAS, 32610 Seri Iskandar, Perak, Malaysia, md_24002668@utp.edu.my
2. Department of Applied Science, Universiti Teknologi PETRONAS, 32610 Seri Iskandar, Perak, Malaysia, mohana.muthuvalu@utp.edu.my
3. Department of Statistics, Noakhali Science and Technology University, Noakhali-3814, Bangladesh, shohel.stat@nstu.edu.bd

Abstract

High-grade gliomas (HGG) are invasive brain tumours characterized by abnormal growth patterns and poor prognoses. Aware and precise prediction of tumour growth helps improve both treatment protocols and patient medical outcome. The quick replicating and diverse nature of HGGs in children makes their predictive progression highly difficult to determine. This research utilized a modified exponential regression approach to study glioma progression in children's brains with predicted accuracy reaching 73.68% for all tumours but elevating to 77.7% for small tumours under 100 mm³. Statistical analyses revealed significant negative correlations between tumour growth and tumour size, along with pre-radiotherapy performance status (PS Before RT), as determined by Kendall's Tau test. The Mann-Whitney U and Kruskal-Wallis H tests were employed for bivariate analysis of categorical data, demonstrating a significant association ($p < 0.05$) among tumour growth rate, the extent of surgical resection, and survival status. The child's age, the occurrence of headaches, and edema were independently associated with the progression of tumour growth. These findings enhance the understanding of paediatric HGGs development, facilitating more accurate prognostic evaluations and improving personalized treatment strategies.

Key Words: High-grade gliomas; Paediatric; Modified exponential regression model; Glioma progression; Prediction.

Mathematical Subject Classification: 62J02, 62P10, 92C50

1. Introduction

Glioma is a malignant brain tumour distinguished by its aggressiveness and poor prognoses (Ma *et al*., 2020). These tumours arise from glial cells, the supporting cells in the brain and nervous system, and are extremely difficult to treat due to their rapid growth and invasive nature. They infiltrate surrounding brain tissue, rendering complete surgical removal difficult and increasing the likelihood of recurrence. Understanding the evolving dynamics of glioma proliferation, particularly the factors influencing tumour growth, is essential for developing effective treatment strategies and improving patient outcomes.

Paediatric gliomas are among the most common and severe types of cancer in children, significantly contributing to cancer related mortality and disability. High-grade gliomas (HGGs) are especially aggressive, posing major challenges to both treatment and long-term survival. HGGs account for approximately $8 - 12\%$ of all paediatric brain tumours, with an average incidence of about 0.85 cases per 100,000 children annually (Ostrom *et al.*, 2022). The occurrence of

these tumours mainly affects children of older age ranges who are between ages 10 and 14, especially seen in this adolescent group. Brainstem tumours called diffuse midline gliomas frequently develop within the brainstem tissue and affect younger patients especially while presenting as DIPGs (Louis *et al.*, 2016). Paediatric HGGs have an unfavourable survival outlook because the 5-year survival chance amounts to only $10-20\%$ and patients with DIPGs have a median survival duration of 9–11 months and often survive less than 10% beyond two years (Jones *et al*., 2017).

Olsson *et al*. (2024) explored the long-term cognitive impacts of paediatric brain tumours, emphasizing the importance of considering cognitive development when evaluating outcomes for children with HGGs. Treatment efforts are complicated by tumour resistance and the potentials of brain damage. Consequently, accurate predictive models and reliable indicators are critical for assessing disease progression, treatment selection, and survival prediction. Despite advancements, there remains a pressing need for improved tools to aid clinical decision-making and personalize treatment (Obeagu & Obeagu, 2024). Statistical models have therefore become essential for enhancing prognostic accuracy and understanding tumour growth in HGG patients. These models support clinicians in risk stratification, personalized treatment planning, and trend forecasting.

Existing literature has predominantly utilized linear and hierarchical regression models to study HGG progression. For instance, Rees *et al.* (2009) used semi-automated segmentation together with hierarchical regression to track tumour volume patterns through which they proved tumour volume functions as the primary indicator of cancer development. Their research model demonstrated that the brain tumour volume grew by 56% each year from six months before transformation. A study utilized linear mixed-effects modelling to study biomarker impact on aggressive HGG growth rates and determined 63.4 days doubling period and 51.6 mm annual diameter expansion (Fan *et al.*, 2020). The prognostic analysis through Cox regression has found four main risk factors that harm patient outlook in paediatric HGGs including pathological grade IV status, thalamic invasion, poor performance status alongside tumours larger than 3.3 cm (Muhammed *et al.,* 2019). The application of multivariate logistic regression predicted post-radiation endocrine dysfunctions in children with successful discrimination power ($AUC = 0.883$) using a clinical threshold of 0.3 (Hua *et al.,* 2012). The use of radiomics in logistic regression models provides excellent accuracy for preoperative glioma grading assessment which enables clinical decision tools for healthcare professionals (Sun *et al.*, 2021).

Despite these advances, current regression frameworks have limited capacity to detect the precise growth patterns of paediatric HGGs, primarily due to tumours rapid and nonlinear proliferation characteristics. To address this limitation, Otunuga (2024) introduced tumour population dynamical methods that incorporate shape parameters, allowing flexible representation of diverse tumours growth trajectories. Modified exponential curves have also been employed to describe and predict the nonlinear progression of real-world phenomena (Lu *et al*., 2012; Zuo *et al*., 2022) and studies suggest that certain tumours, including glioma, follow modified exponential growth patterns during development. Feucht *et al.* (2024) observed that glioblastomas often exhibit exponential growth in their early stages, with a median volume doubling period of 31 days as confirmed by an exponential regression model. The utilization of baseline tumour volume together with heterogeneous clinical data remains inconsistent for improving prognostic accuracy.

Dynamics of HGG and development are not always linear; in most cases it has an exponential or sigmoid asymptote due to their rapid proliferation. Where the rate of change accelerates, such as is the case in malignant tumours, the linear regression approach fails to characterize this accelerating growth. Logistic and Cox regression models can be used in classification and survival analysis models but cannot be used in nonlinear modelling growth trajectories. Exponential regression (Mahanty *et al.*, 2022), in contrast, more closely corresponds to the biological reality of nonlinear glioma growth because each cell division leads to tumour expansion in a multiplicative approach. Clinically, this gives better prediction of tumour burden outcome which can be used in the timely planning of treatment.

This study presents a generalized exponential regression model tailored for paediatric high-grade gliomas, adept at identifying both rapid growth and decrease patterns that traditional linear regression models overlook. This approach combines tumour growth evaluation with clinical and patient-specific factors. These coefficients can aid clinical teams in optimizing the timing of surgical resections and treatment measures. The model provides physicians with an interactive instrument for enhancing the accuracy of intervention decisions, potentially reducing developmental difficulties in paediatric patients.

2. Data Resource

This study utilizes open-source data acquired from the Kaggle database (2024)**.** This collection contains clinical and tumour size data for 57 paediatric cancer patients aged 2 to 17 years. The tumour volumes of each patient will be recorded as v_i, indicating the tumour size, followed by the calculation of the percentage growth rate (PGR) to determine if there is rapid growth or decline, using the following formula (Casler, 2015)

$$PGR = \frac{V_{i+1}-V_i}{V_i} \times 100 \quad (1)$$

where, v_i is the i^{th} tumour volume in mm^3 and v_{i+1} is the $(i+1)^{th}$ tumour volume in mm^3. If PGR is positive, it indicates rapid growth, whereas a negative PGR suggests a decline in growth. To enhance model accuracy, PGR is utilized as a continuous variable for fitting the model, while also being classified into categorical groups (rapid growth and decline) to examine its relationship with other factors. However, the outcome variable violated the linearity assumption concerning initial tumour volume ($r = -0.29$), indicating a nonlinear relationship. Clinical characteristics of paediatric patients are shown in Table 1.

Table 1: Clinical characteristics of HGG patients for categorical variables.

Characteristics	Categories	Number (%)
Growth	Rapid Growth	26 (45.6)
	Decline	31 (54.4)
Gender	Male	31 (54.4)
	Female	26 (45.6)
Headache	Yes	40 (70.2)
	No	17 (29.8)
Edema	Mild	23 (40.3)
	Moderate	29 (50.9)
	No	1 (1.8)
	Severe	4 (7.0)
Neurostate	Asymptom	17 (29.8)
	Symptoma	40 (70.2)
Extent of Surgical resection	Biopsy	22 (38.6)
	Gross Resection	10 (17.5)
	Subtotal	25 (43.9)

3. Statistical Analysis

This study included qualitative and quantitative statistical methods. Initially, we utilized the Kolmogorov-Smirnov test to evaluate if the result variable adhered to a normal distribution. Kendall's Tau was utilized to assess the nonparametric correlations among growth, volume, age, and PS before RT. Furthermore, the non-parametric Mann-Whitney U test and the Kruskal-Wallis H test were employed to examine the associations among categorical variables. The Mann-Whitney U test was utilized when the result variable exhibited a non-normal distribution, and the independent variables were categorical with two categories. Likewise, the Kruskal-Wallis H test was employed under identical conditions for qualitative independent variables with three or more categories. A two-tailed probability with a significance level of 0.05 was uniformly employed. Secondly, a correlation test and a normal probability-probability plot of residuals were employed to confirm that the growth rate and volume demonstrate a nonlinear trend. Nonlinear models effectively represent the diverse connection between tumour volume and growth rate, particularly when the data violates linearity assumptions. A generalized exponential regression model was utilized to forecast glioma development in nonlinear progressions. Let us consider,

$$y = a + bv + kc^v + e \quad (2)$$

where, y is the percentage growth rate (PGR), v is the volume and a, b, c and k are constant parameters such that, $-\infty < a, b < \infty$, $c > 0$, $k \neq 0$ and e is the statistical error term, assumed to be normally distributed with mean zero and constant variance. The best model fit is described in Table 2 and coefficient of determination $(R^2) = 0.775$. The R^2 indicates that 77.5% of the variation in the dependent variable growth of paediatric is explained by the independent variable initial tumour volume.

Table 2: Parameter estimation

			95% confidence interval	
Parameter	**Estimate**	**SD**	**Lower bound**	**Upper bound**
a	78.307	154.380	-231.330	387.950
b	-0.088	0.182	-0.454	0.277
c	0.962	0.007	0.948	0.975
k	7852.010	769.360	6308.870	9395.160

For the diagnostic tests, the runs test was applied to evaluate the randomness of residuals, ensuring that the residuals exhibited no systematic patterns. Additionally, a spread-location plot, also known as a scale-location plot, was utilized to examine the homoscedasticity of residuals verifying that residuals maintain constant variance across all levels of the independent variable. All statistical analyses were performed using SPSS software version 15.0.

4. Results and Discussion

4.1 Normality test

Table 3: Test for normality of the paediatric glioma growth.

	Kolmogorov-Smirnov		
	Statistic	**df**	**Sig.**
HGG Growth	0.387	57	0.000

The normality of the outcome variable was assessed using the Kolmogorov-Smirnov test, which confirmed the non-normality assumption ($p < 0.05$) as in Table 3 for the study variables.

4.2 Nonparametric Association for non-normal numeric: Kendall's Tau

Table 4: Nonparametric correlation analysis of growth, volume, PS before RT and age.

	HGG Growth	**Volume (mm^3)**	**PS before RT**	**Age (years)**
HGG Growth	1.000			
Volume	**-0.415****	1.000		
PS before RT	0.188	**-0.262****	1.000	
Age	0.086	0.006	-0.066	1.000

**Correlation is significant at the 0.05 level (2-tailed)

Nonparametric Kendall's Tau test indicates a significant negative relationship between growth rate and volume ($r = -0.415$), and between volume and PS prior to RT ($r = -0.262$), both significant at the 0.05 level. Increasing the volume and PS before RT leads to a decreased growth rate, signifying a significant inverse correlation as shown in Table 4. Extremely weak association exists between the development of glioma and child age.

4.3 Association Analysis for qualitative independent variables: Mann-Whitney U test and Kruskal–Wallis H test

Table 5: Test of significant mean difference of growth between the categorical variables.

Variables	**Mann-Whitney U test**	**Kruskal–Wallis H test**	**P value**	**Remark**
Survival status (Dead, Alive)	185.0	-	**0.003**	Significant
Gender (Male, Female)	399.0	-	0.949	Not Significant
Headache (Yes, No)	268.0	-	0.209	Not Significant
Extent of Surgical resection (Biopsy, Gross resection, Subtotal)	-	7.675	**0.022**	Significant
Edema (Mild, Moderate, No, Severe)	-	6.116	0.106	Not Significant

The Mann- Whitney U test in Table 5 showed no significant difference in growth between male and female (Mann-Whitney U= $399.0, p > 0.05$), and Headache (Mann-Whitney U= $268.0, p > 0.05$) and Edema (Kruskal-Wallis H= $6.116, p > 0.05$). However, Survival status was found to be significant with growth of children (Mann-Whitney U= $185.0, p < 0.05$). From Table 5, it is also observed that extent of surgical resection (Kruskal-Wallis H= $7.675, p < 0.05$) was significant with glioma growth.

4.4 Quantitative analysis

Table 6. Basic statistical characteristics of age, gross tumour volume and PS Before RT for continues variables.

Variables	Mean	Median	SD	Range	Skewness	Kurtosis
Age (years)	7.99	8.00	3.98	15	0.428	-0.767
Volume (mm^3)	588.65	511.28	520.80	2247.7	1.405	1.877
PS Before RT	56.49	60.00	17.16	50	-0.109	-1.310

According to Table 6, the mean age of the patient demographics is 7.99 years, with a median age of 8 years and a range of 15 years. The average volume of children is 588.65 mm^3, while the median PS before to RT is 60, indicating a need for occasional support in caring for children with glioma cancer.

4.5 Pearson correlation test and linearity test

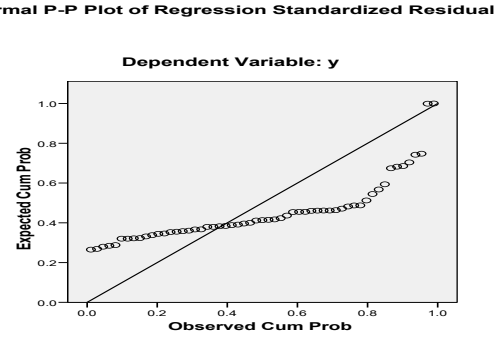
Pearson correlation	Normal p-p plot of residual for linear model
-0.29	

Figure 1: Pearson correlation and linearity test

The results demonstrate a low association between the growth rate and volume, indicating that these variables are not significantly linearly associated. The normal probability-probability (P-P) plot in Figure 1, indicating that the residuals of the linear model are not randomly distributed, thereby violating the randomness assumption necessary for linear regression. The findings validate that the provided data demonstrates nonlinear behaviour. Thus, a nonlinear generalized exponential regression model will be utilized to more precisely forecast outcomes for paediatric glioma patients, since it is more adept at capturing the intricate correlations present in the data.

4.6 Prediction and estimation of paediatric HGG glioma growth

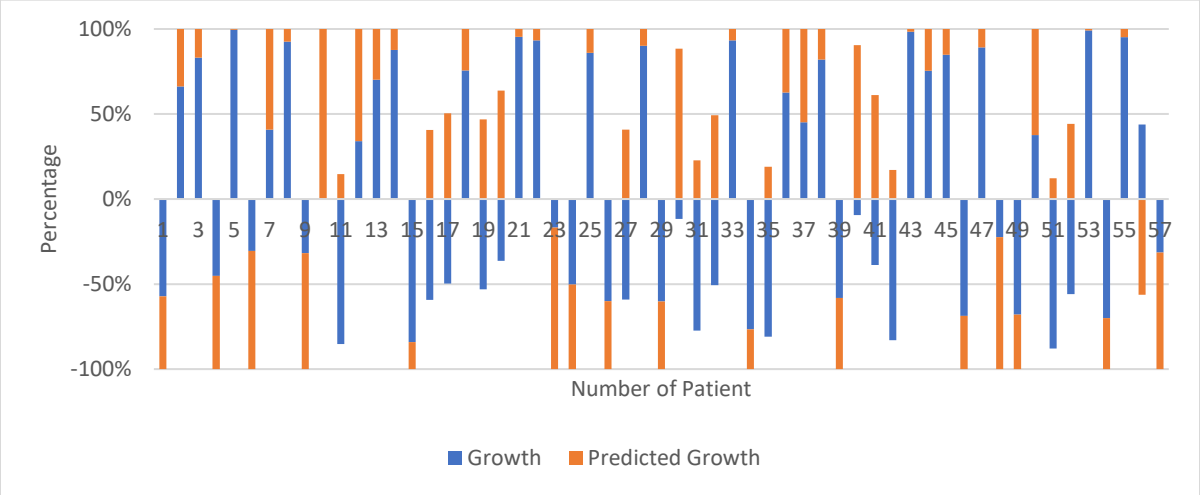


Figure 2: Prediction of glioma growth for rapid growth and decline.

A high percentage denotes rapid growth, whereas a negative percentage implies a decrease in growth rates as shown in Figure 2. The simultaneous movement of the two colours (blue and orange) indicates that the model effectively predicts glioma proliferation. Conversely, opposing trajectories of the two colours indicate suboptimal predicting performance. The findings provide significant insights that can enhance early diagnosis, risk evaluation, and the formulation of individualized treatment regimens for paediatric high-grade glioma patients.

Table 7: Prediction of the paediatric glioma growth from the fitted model

	No. of Patients	No. of Patients Predicted	Correct Prediction (in %)
HGG volume (in mm^3)	57	42	73.68%
HGG volume < 100 mm^3	9	7	77.70%

The model has a 73.68% accuracy in predicting glioma development within the chosen dataset, accurately detecting 42 of 57 paediatric cases. It attained a superior accuracy rate of 77.70% for small glioma tumours, successfully predicting 7 out of 9 instances with volumes less than $100\ mm^3$. This indicates that the model is very proficient at predicting the growth of small tumours, surpassing its general prediction accuracy as illustrated in Table 7. The results underscore the model's efficacy as a dependable instrument for the early diagnosis and surveillance of paediatric glioma, particularly in instances with lower tumour sizes.

4.7 Test for residuals: Run test

Table 8: Run test of residuals

Test value	Number of Runs	Sig.(2-tailed)/p value
-30.270	26	0.351

Table 8 describes the diagnostic assessment of residuals using the run test. Since $p > 0.05$, the residuals derived from the calculated model are random. The results confirmed that the fitted model accurately predicts the HGG growth of children.

4.8 Spread-Location Plot

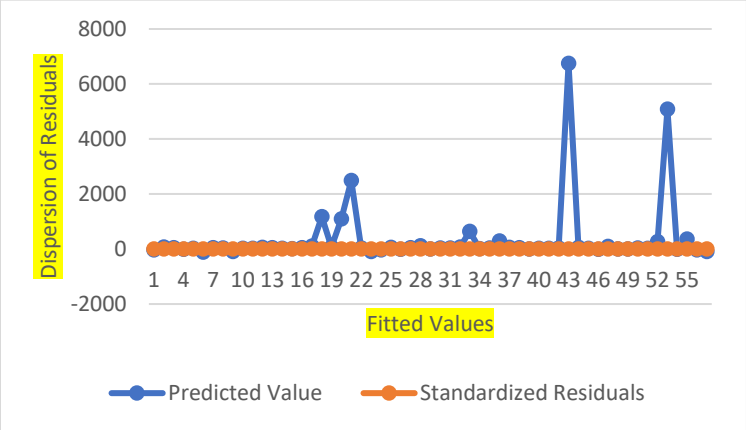


Figure 3: Scale-Location plot for residuals

This figure 3 assists the evaluation of homoscedasticity (constant variance of residuals) in regression models by illustrating the association between fitted values and the dispersion of residuals. The diagram indicates that the residuals exhibit constant variance. Given that the average errors ($\sum e_i = 0$) is zero and the variance is constant, the fitted model accurately predicts HGG in paediatric patients.

5. Conclusion

The study employs a generalized exponential regression model to predict growth patterns in paediatric gliomas, achieving an accuracy rate of 73.68%. This model is particularly significant for paediatric patients with HGGs exhibiting nonlinear behaviour. The findings reveal a strong correlation between the growth rate of paediatric HGGs and two critical factors: the extent of surgical resection and the survival status (alive or deceased). It suggests improving early diagnosis, risk assessment, and the development of personalized treatment strategies for children with glioma. However, certain factors, such as the patient's functional level and immunological symptoms, which may influence tumour development patterns, were not evaluated in this study. Despite these limitations, this research introduces a novel and effective approach to estimating glioma progression in paediatric patients under diverse conditions. Future studies should focus on validating the model using larger datasets and exploring its potential to enhance clinical decision-making and patient outcomes.

Data Availability Statement

The data supporting this investigation are accessible from the Kaggle database. Data are available via the Kaggle Data Hub at https://www.kaggle.com/datasets/amraam/pediatric-high-grade-glioma-dataset/dat (accessed December 2024) (Kaggle database, 2024).

Funding Statement

This research was funded by Ministry of Higher Education, Malaysia under the Fundamental Research Grant Scheme (Cost Center: 0153AB-L52).

Ethics Approval

Due to the dataset being publicly accessible and de-identified, institutional review board (IRB) approval and additional patient consent were not required for this secondary study.

References

1. Casler, S. D. (2015). Why Growth Rates? Which Growth Rate? Specification and Measurement Issues in Estimating Elasticity Values. *The American Economist*, *60*(2), 142-161.
2. Fan, Z., Liu, Y., Li, S., Liu, X., Jiang, T., Wang, Y., & Wang, L. (2020). Association of tumor growth rates with molecular biomarker status: a longitudinal study of high-grade glioma. *Aging (Albany NY)*, *12*(9), 7908-7926.

3. Feucht, D., Haas, P., Skardelly, M., Behling, F., Rieger, D., Bombach, P., Paulsen, F., Hoffmann, E., Hauser, T. K., Bender, B., Renovanz, M., Niyazi, M., Tabatabai, G., Tatagiba, M., & Roder, C. (2024). Preoperative growth dynamics of untreated glioblastoma: Description of an exponential growth type, correlating factors, and association with postoperative survival. *Neuro-Oncology Advances*, *6*(1), 1-12.
4. Hua, C., Wu, S., Chemaitilly, W., Lukose, R. C., & Merchant, T. E. (2012). Predicting the probability of abnormal stimulated growth hormone response in children after radiotherapy for brain tumors. *International Journal of Radiation Oncology Biology Physics*, *84*(4), 990-995.
5. Jones, C., Karajannis, M. A., Jones, D. T., Kieran, M. W., Monje, M., Baker, S. J., Becher, O. J., Cho, Y.-J., Gupta, N., & Hawkins, C. (2017). Pediatric high-grade glioma: biologically and clinically in need of new thinking. *Neuro-Oncology*, *19*(2), 153-161.
6. Kaggle database (2024). *Pediatric High grade Glioma Dataset*. Kaggle. https://www.kaggle.com/datasets/amraam/pediatric-high-grade-glioma-dataset/data
7. Louis, D. N., Perry, A., Reifenberger, G., Von Deimling, A., Figarella-Branger, D., Cavenee, W. K., Ohgaki, H., Wiestler, O. D., Kleihues, P., & Ellison, D. W. (2016). The 2016 World Health Organization classification of tumors of the central nervous system: a summary. *Acta Neuropathologica*, *131*(6), 803-820.
8. Lu, Y., Li, Q., & Guo, Y. (2012). Improved prediction model of modified exponential curve. In *2012 Third International Conference on Intelligent Control and Information Processing* (pp. 513-516). IEEE.
9. Ma, Z., Niu, B., Phan, T. A., Stensjøen, A. L., Ene, C., Woodiwiss, T., Wang, T., Maini, P. K., Holland, E. C., & Tian, J. P. (2020). Stochastic growth pattern of untreated human glioblastomas predicts the survival time for patients. *Scientific Reports*, *10*(1), 6642, 1-13.
10. Mahanty, C., Kumar, R., Mishra, B. K., Hemanth, D. J., Gupta, D., & Khanna, A. (2022). Prediction of COVID-19 active cases using exponential and non-linear growth models. *Expert Systems*, *39*(3), e12648, 1-22.
11. Muhammed, A., Gaber, M. S., Elbeltagy, M., El Hemaly, A., Taha, H., Refaat, A., & Zaghluol, M. S. (2019). Risk stratification of pediatric high-grade glioma: a newly proposed prognostic score. *Child's Nervous System*, *35*(12), 2355-2362.
12. Obeagu, E. I., & Obeagu, G. U. (2024). Predictive models and biomarkers for survival in stage III breast cancer: a review of clinical applications and future directions. *Annals of Medicine and Surgery*, *86*(10), 5980-5987.
13. Ostrom, Q. T., Price, M., Neff, C., Cioffi, G., Waite, K. A., Kruchko, C., & Barnholtz-Sloan, J. S. (2022). CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2015–2019. *Neuro-Oncology*, *24*(Supplement_5), v1-v95.
14. Otunuga, O. M. (2024). Tumor growth and population modeling in a toxicant-stressed random environment. *Journal of Mathematical Biology*, *88*(2), Article18.
15. Rees, J., Watt, H., Jäger, H. R., Benton, C., Tozer, D., Tofts, P., & Waldman, A. (2009). Volumes and growth rates of untreated adult low-grade gliomas indicate risk of early malignant transformation. *European Journal of Radiology*, *72*(1), 54-64.
16. Sun, X., Liao, W., Cao, D., Zhao, Y., Zhou, G., Wang, D., & Mao, Y. (2021). A logistic regression model for prediction of glioma grading based on radiomics. *Journal of Central South University. Medical Sciences*, *46*(4), 385-392.
17. Tonning Olsson, I., Lundgren, J., Hjorth, L., Munck Af Rosenschöld, P., Hammar, Å., & Perrin, S. (2024). Neurocognitive development after pediatric brain tumor-a longitudinal, retrospective cohort study. *Child Neuropsychology*, *30*(1), 22-44.
18. Zuo, K., Zuo, H., & Zhang, X. (2022). An Exponential Curve Model and Its Application in Forecasting Private Car Ownership of China. *Journal of Mathematics*, *2022*(1), 1-6.