ISSN 2282-0930 • Epidemiology Biostatistics and Public Health - XIII^th SISMEC Congress - Vol. 20 Suppl. 1

Integrative Deep Learning of Germline and Somatic Genomics in Glioblastoma: A Translational Approach to Prognosis and Drug Target Discovery

Polano M.[1], Ius T.[2,3], Di Loreto C.[3], Toffoli G.[1], Dal Bo M.[1], Lombardi G.[4], Miran Skrap[3]

(1) Experimental and Clinical Pharmacology Unit, Centro di Riferimento Oncologico di Aviano (CRO) IRCCS, Via Franco Gallini, 2, 33081, Aviano, Italy
(2) Academic Neurosurgery, Department of Neurosurgery, University of Padova, 35121 Padova, Italy
(3) Neurosurgery Unit, Head-Neck and Neuroscience Department, University of Udine 33100, Udine, Italy
(4) Department of Medicine, University of Udine, Udine, University Hospital of Udine, Piazzale Santa Maria della Misericordia 15, 33100 Udine, Italy
(5) Department of Oncology 1, Veneto Institute of Oncology IOV-IRCCS, Padua, Italy

CORRESPONDING AUTHOR: *Polano M., mpolano@cro.it*

INTRODUCTION

Glioblastoma (GBM) is the most aggressive primary malignant brain tumor in adults, characterized by rapid progression, high recurrence rates, and resistance to conventional therapies. The standard-of-care treatment, based on maximal safe surgical resection followed by radiotherapy and concomitant/adjuvant temozolomide (Stupp protocol), has modestly improved patient outcomes since 2005. However, the median overall survival (OS) remains limited to approximately 12–16 months, and progression-free survival (PFS) to 6–10 months. A major clinical challenge is the heterogeneity in patient response, which is partially attributable to underlying genetic factors—both germline and somatic. While somatic drivers of GBM have been extensively studied, the role of germline variation and its interaction with somatic alterations in modulating treatment response and survival outcomes remains underexplored.

AIM

The aim is to improve risk stratification in glioblastoma by developing and validating an interpretable machine learning model that integrates germline and somatic genomic alterations with clinical variables, particularly extent of resection (EOR), to predict progression-free and overall survival. The model aims to:

- enhance the early identification of patients at higher risk of rapid progression or death; and
- uncover molecular signatures—across germline variants, somatic mutations, and gene amplifications—that are indicative of a poor prognosis and may inform more personalised therapeutic strategies.

MATERIALS AND METHODS

We retrospectively analyzed a cohort of 119 patients diagnosed with primary GBM who underwent surgical resection at the Department of Neurosurgery, Udine Hospital, between 2014 and 2019. Inclusion criteria were availability of high-quality tumor and blood samples, comprehensive clinical data, and follow-up. All patients received the Stupp protocol, and 44 patients (38%) were additionally treated with carmustine wafer implantation. Targeted next-generation sequencing (NGS) was performed using a multi-gene panel to profile both germline single nucleotide polymorphisms (SNPs) and somatic variants, including point mutations and gene amplifications. After quality filtering, 1,192 high-confidence germline SNPs were retained from an initial 4,680 variants. Somatic alterations in key oncogenes and tumor suppressors (e.g., EGFR, MDM2, TP53, PDGFRA) were encoded as binary features. We employed the GenNet framework, an interpretable deep learning architecture tailored for genotype-phenotype prediction, to model PFS and OS as continuous outcomes. The model incorporated EOR and age as covariates and was trained using mean squared error (MSE) loss. Model performance was benchmarked against traditional machine learning methods, including Random Survival Forests (RSF), to assess prediction accuracy and feature interpretability.

DOI: 10.54103/2282-0930/29479

RESULTS

The median overall survival (OS) in the cohort was 16 months (95% CI: 15–19), and the median progression-free survival (PFS) was 10 months (95% CI: 8–12). The median patient age was 60 years (95% CI: 52–69), and the median extent of resection was 98% (95% CI: 95–100).

The GenNet model achieved a test-set MSE of 63.93 for survival prediction (batch size = 32, learning rate = 0.001, epochs = 200, L1 regularization = 0.01). Key predictive features included both germline and somatic variants in genes such as TP53, PAX7, PIK3C2G, CYLD, FGF5, ERG, PIK3R2, and ALK. The inclusion of somatic gene amplifications notably enhanced the model's accuracy and biological interpretability. In comparative analysis, GenNet outperformed traditional RSF models in both predictive power and feature attribution.

CONCLUSIONS

Our study demonstrates the value of integrating germline and somatic genomic information with clinical variables in predicting survival outcomes in glioblastoma. The interpretable deep learning model built using the GenNet framework provides insight into the genomic architecture of disease progression and holds potential for informing personalized therapeutic strategies. These findings underscore the relevance of multi-omic modeling approaches in precision neuro-oncology and support further validation in larger, prospective cohorts.

DOI: 10.54103/2282-0930/29479