Clinical Significance and Molecular Annotation of Cellular Morphometric Subtypes in Lower Grade Gliomas discovered by Machine Learning: a retrospective multicentric study

Xiao-Ping Liu

Lawrence Berkeley National Lab

Xiaoqing Jin

Zhongnan Hospital of Wuhan University

Saman Ahmadian

Stanford University Medical Center <https://orcid.org/0000-0002-4133-7895>

Xu Yang

Nanjing Medical University

Su-Fang Tian

Zhongnan Hospital of Wuhan University

Yu-Xiang Cai

Zhongnan Hospital of Wuhan University

Kuldeep Chawla

Lawrence Berkeley National Lab

Antoine Snijders

Lawrence Berkeley National Laboratory <https://orcid.org/0000-0002-7086-1063>

Yankai Xia

Nanjing Medical University

Paul van Diest

University Medical Center Utrecht

Jian-Hua Mao

Lawrence Berkeley National Laboratory <https://orcid.org/0000-0001-9320-6021>

Zhi-Qiang Li

Zhongnan Hospital of Wuhan University

Hannes Vogel

Stanford University Medical Center

Hang Chang (✉ HChang@lbl.gov)

Lawrence Berkeley National Lab <https://orcid.org/0000-0002-3773-6818>

Article

Keywords: lower grade glioma, Glioblastoma, Stacked Predictive Sparse Decomposition, cellular morphometric biomarkers, cellular morphometric subtypes, overall survival, nomogram, immunohistochemistry staining

Posted Date: August 12th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-770415/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Abstract

Lower grade gliomas (LGGs) are heterogeneous diseases by clinical, histological and molecular criteria. Here, we developed a machine learning pipeline to extract cellular morphometric biomarkers from whole slide images of tissue histology; and identified and externally validated robust cellular morphometric subtypes of LGGs in multi-center cohorts. The subtypes have significantly independent predictive power for overall survival across all three independent cohorts. In the TCGA-LGG cohort, we found that patients within the poor-prognosis subtype responded poorly to primary therapy and follow-up treatment. Furthermore, LGGs within the poor-prognosis subtype were characterized by higher mutational burden, higher frequencies of copy number alterations, and higher level of tumor-infiltrating lymphocytes and immune checkpoint genes. Higher level of PD-1/PD-L1/CTLA-4 was confirmed by immunohistochemical staining. In addition, the subtypes learned from LGG demonstrates translational impact on Glioblastoma (GBM). Overall, we developed and validated a framework for the cellular morphometric subtype discovery in LGGs associated with specific molecular alterations, immune micro-environment, prognosis and treatment response.

Introduction

Gliomas are the most common primary central nervous system (CNS) malignant tumor. According to statistics from the Central Brain Tumor Registry of the United States (CBTRUS), gliomas account for about 80% of all central nervous system malignancies^{1,2}. Historically, according to the 2007 WHO classification criteria for CNS tumors, gliomas were categorized into grade 1, 2, 3 or 4³. With recent integration of molecular alterations in the classification of CNS tumor, the newest WHO classification (2016) introduced a paradigm shift in the classification of CNS tumors combining both histopathologic and genotypic features⁴ to reveal an “integrated” diagnosis that better correlates with patient outcome. LGGs grow relatively slowly, and due to their characteristic diffuse and infiltrative growth, they are prone to recurrence after surgical resection, occasionally accompanied by an increase in malignancy grade, and have a high clinical disability and mortality rate⁵. In contrast, glioblastoma multiforme (GBM, grade 4) displays a much more unfavorable course.

It is reported that age > 40 years, astrocytic subtype, tumor maximum diameter > 6 cm, tumors crossing the midline, and the patient’s degree of neurological impairment are risk factors affecting overall survival (OS)^{6,7}. For patients with low risk LGG, the median OS can reach 10.8 years, while for patients with high risk LGG (2 or more aforementioned risk factors), the median OS is only 3.9 years. Moreover, LGGs are highly heterogeneous both at the histopathological and molecular level^{4,8}, reflected in significant variability in clinical outcome^{8,9}. Therefore, to personalize care and treatment of LGG patients, accurate and robust patient stratification, which is significantly associated with clinical outcomes, is mandatory.

It is well known that cellular morphometric properties play a key role in cancer diagnosis and cancer prognosis together with important molecular factors. Although deep neural networks (e.g., CNN-like

models) have been successfully applied in many glioma-related studies, including diagnosis¹⁰⁻¹², grade prediction^{13,14}, and mutation prediction^{15,16}, based on computed tomography (CT), magnetic resonance imaging (MRI), and whole slide images (WSI), the quantitative profiling and molecular association of cellular morphometric landscape from whole slide images is inadequately investigated due to both technical and conceptual limitations.

To capture the heterogeneous cytoarchitecture of gliomas, we developed an high-throughput computational pipeline that enables quantification of tissue histology at the cellular level, insensitive to variations in fixation and staining protocols at different laboratories¹⁷. The extracted cellular morphometric features can be used for tumor classification¹⁸, and molecular association¹⁹. In addition, our previous study introduced stacked predictive sparse decomposition (SPSD)²⁰ that could be used to mine underlying cellular morphometric properties within whole slide images (WSI). In the present study, we applied SPSPD to LGG cohorts to discover clinically relevant cellular morphometric subtypes (CMSs), and evaluated clinical impacts and molecular correlation of CMSs.

Results

Study Design and Characteristics of Patient Cohorts

We used three retrospective cohorts of LGG patients to evaluate and independently validate the prognostic impact of cellular morphometric subtypes, which were identified from WSIs powered by artificial intelligence. And we also used two retrospective cohorts of GBM patients to evaluate the generalization capability and translational impact of cellular morphometric subtypes learned from LGG onto GBM (Fig. 1). Specifically, the TCGA-LGG cohort served as the discovery set and included a total of 488 LGG patients with both diagnostic histopathological slides and clinical information. There were 271 (55.5%) male and 217 (44.5%) female patients, with a median age of 41 years (range: 14-87 years) (Supplementary Table 1). The ZN-LGG cohort included 70 LGG patients, where 36 (51.4%) patients were male and 34 (48.6%) patients were female, and the median age were 47.0 years (range: 6-72 years) (Supplementary Table 2). The SU-LGG cohort included 37 LGG patients, where 22 (59.5%) patients were male and 15 (40.5%) patients were female, and the median age was 41.0 years (range: 1-83 years) (Supplementary Table 3). The TCGA-GBM cohort included 380 GBM patients, where 145 (38.2%) patients were male and 234 were (61.6%) patients were female, with a median age of 59.0 years (range: 10-89 years) (Supplementary Table 4). The ZN-GBM cohort included 79 LGG patients, where 23 (29.1%) patients were male and 55 (69.6%) patients were female, and the median age were 56.0 years (range: 5-81 years) (Supplementary Table 5).

Identification of Cellular Morphometric Biomarkers using Unsupervised Representation Learning

In this study, our cellular delineation pipeline¹⁷ recognized and delineated over 400 million cellular objects from 841 diagnostic slides of 488 TCGA-LGG patients; over 25 million cellular objects from 75 diagnostic slides of 70 ZN-LGG patients; over 10 million cellular objects from 37 diagnostic slides of 37

SU-LGG patients; over 400 million cellular objects from 848 diagnostic slides of 380 TCGA-GBM patients; and over 25 million cellular objects from 85 diagnostic slides of 79 ZN-GBM patients, where each cellular object was represented with 15 morphometric properties (Extended Data Fig. 1a and Supplementary Table 6).

Next, we optimized and trained our SPSPD²⁰ model based on pre-quantified cellular objects randomly selected from TCGA-LGG cohort to discover the underlying cellular morphometric biomarkers (Extended Data Fig. 2). After training, the pre-built SPSPD model reconstructed each cellular object as a sparse combination of the pre-identified 256 cellular morphometric biomarkers, which led to the novel representation of each single cellular object as the 256 sparse code (reconstruction coefficient); and thereafter, the corresponding 256-dimensional cellular morphometric context representation of each patient as an aggregation (Extended Data Fig. 1b) of all delineated cellular objects belonging to the same patient (Supplementary Tables 7-11). The final patient-level cellular morphometric context representation was optimized by using the top 30 (sparsity constraint of SPSPD model) out of 256 cellular morphometric biomarkers with the largest variations, which contributed to 98.84% of the total variations of the data.

Clinical and Biological Evaluation of Cellular Morphometric Biomarkers

We next evaluated association of the 30 cellular morphometric biomarkers with respect to histological meanings, prognosis as well as cancer biology. Our survival analysis revealed that 20 out of 30 cellular morphometric biomarkers had significant prognostic impact ($FDR < 0.05$), where 5 of them were prognostically favorable (hazard ratio (HR) < 1) and 15 prognostically unfavorable (HR > 1) (Fig. 2a, Extended Data Fig. 3 and Supplementary Table 12). Examples of prognostically significant cellular morphometric biomarkers (Fig. 2a, and Extended Data Fig. 3) demonstrated the capability of our unsupervised learning pipeline in acquiring biomedically meaningful and interpretable histopathological concept at the cellular level.

Additionally, the TCGA-LGG patient cohort was divided into two groups based on each cellular morphometric biomarker. The Kaplan-Meier curves showed significant impact ($P < 0.01$, Fig. 2b, and Extended Data Fig. 4) of the levels of each cellular morphometric biomarker on OS. Thereafter, we evaluated biological significance between patient groups with high and low cellular morphometric biomarker levels in the TCGA-LGG cohort and discovered significant correlations ($P < 0.05$) with tumor microenvironment factors, including tumor immune cells and fibroblast²¹, and predictors of immunotherapy response (Fig. 2c, and Extended Data Fig. 5 and 6). Levels of prognostic favorable cellular morphometric biomarkers correlated negatively, while levels of prognostic unfavorable cellular morphometric biomarkers correlated positively with the number of tumor-infiltrating immune cells and the expression levels of PD-1 and PD-L1, but not to the number of fibroblasts ($P > 0.05$; Fig. 2c, and Extended Data Fig. 5 and 6). Finally, we detected a significant correlation with SCNA and TMB ($P < 0.05$; Fig. 2d).

Identification and Validation of Cellular Morphometric Subtypes

Consensus cluster analysis using 30 cellular morphometric biomarkers identified three subtypes from TCGA-LGG cohort with a significantly different prognosis (Log-rank $P < 0.0001$; Extended Data Fig. 7). Given the small number of patients ($n=4$) in subtype 3, as well as its prognostic similarity to subtype 2 patients, we merged subtypes 3 and 2, and thereafter refer this combination as subtype 2 in the rest of this study (Fig. 3a). Accordingly, the TCGA-LGG cohort contained 389 subtype 1 and 99 subtype 2 patients. The patient-level cellular morphometric context representation in the TCGA-LGG cohort formed significantly distinct clusters ($P = 0.001$, Fig. 3b). Importantly, two CMSs, predicted with pre-built subtype model, were portioned in two validation sets. Specifically, the ZN-LGG cohort stratified into subtype 1 (38 patients) and subtype 2 (32 patients); while the SU-LGG cohort stratified into subtype 1 (16 patients) and subtype 2 (21 patients). Moreover, the patient-level cellular morphometric context representation in both validation cohorts also formed significantly distinct clusters ($P=0.001$, Fig. 3e, h).

Clinical significance of Cellular Morphometric Subtypes

We examined the association between CMSs and clinical and tumor characteristics in the TCGA-LGG cohort. Surprisingly, there was no significant association between CMSs and any clinical or molecular prognostic factor (including, age, grade, histological type, IDH mutation, 1p/19q codeletion, MGMT promoter status, TERT promoter status and ATRX status) (Supplementary Table 1). This finding was confirmed in both validation cohorts (Supplementary Tables 2 and 3).

In the TCGA-LGG cohort where genetic alteration burden information was available, Maftool analysis showed significantly higher TMB ($P=0.003$) and SCNA score ($P=0.012$) in subtype 2 patients (Extended Data Fig. 8), indicating higher level of genomic instability of tumors from subtype 2.

Kaplan-Meier analysis showed significantly shorter OS of subtype 2 patients than subtype 1 patients ($P=0.001$, Fig. 3c). Furthermore, univariate and multivariate CoxPH models indicated the independent prognostic impact of CMSs in the TCGA-LGG cohort after adjusting for other significant clinical and molecular factors including age, histological type, grade, IDH mutation status, and ATRX mutation status (HR: 1.773, 95% CI: 1.066-2.947, $P = 0.027$; Fig. 3d, and Supplementary Table 13). The combination of CMSs and clinical and molecular factors provided significantly improved ($P < 0.001$, Extended Data Fig. 9a) prediction of OS (median C-index: 0.860, 95% CI: 0.859-0.861) compared to classical models with only clinical and molecular factors (median C-index: 0.857, 95% CI: 0.856-0.858). Moreover, the nomogram (Fig. 4a), built upon patient subtype and clinical and molecular factors, significantly correlated with OS of TCGA-LGG patients, and provided excellent prediction (C-indexes for validation on the training set and test set with 1000 bootstraps were 0.8334 (95% CI: 0.8322- 0.8345) and 0.8014 (95% CI: 0.8001- 0.8026), respectively) of the 3-year and 5-year OS of TCGA-LGG patients, which was further confirmed by calibration analysis on the training (Fig. 4b,c) and testing set (Fig. 4d,e), respectively. In clinical practice, [physicians](#) can draw a vertical line from each variable to the "Point" line (Fig. 4a) to calculate the score of the corresponding variable, and thereafter the total score of a patient as the sum of all individual score above, with which, the clinician can predict the 3-year and 5-year OS rates of a new LGG patient. Meanwhile, a dynamic nomogram further demonstrates its potential clinical implications

at: https://liuxiaoping.shinyapps.io/LGG_nomogram/. In addition, Chi-square test showed significantly poor response of subtype 2 patients with respect to both primary therapy ($P<0.001$) and follow-up treatment ($P=0.002$) (Supplementary Table 1).

Importantly, the double-blind deployment of the pre-built patient subtype model on both validation cohorts followed by independent survival analysis confirmed the significantly worse OS of subtype 2 patients ($P=0.027$ in ZN-LGG, $P=0.005$ in SU-LGG, Fig. 3f, i). Furthermore, univariate and multivariate CoxPH models confirmed the independent prognostic impact of CMSs after adjustment for other significant clinical factors in both validation cohorts (ZN-LGG: HR: 4.776, 95% CI: 1.29-17.686, $P=0.019$; SU-LGG: HR: 9.392, 95% CI: 1.944-45.373, $P=0.005$; Fig. 3g, j, and Supplementary Tables 14 and 15).

Interestingly, the direct translation and deployment of the pre-built patient subtype model on TCGA-GBM and ZN-GBM (double-blind) cohorts followed by independent survival analysis confirmed the clinical impact of subtype learned from LGG on GBM patients (Fig. 5). Consistent with our observations on LGG cohorts, GBM patients in both cohorts were stratified into distinct clusters ($P=0.001$ in TCGA-GBM; $P=0.001$ in ZN-GBM; Fig. 5 a, b), and the subtype 2 GBM patients demonstrated significantly worse OS compared with subtype 1 GBM patients ($P=0.00051$ in TCGA-GBM; $P<0.0001$ in ZN-GBM; Fig. 5c, d). Furthermore, univariate and multivariate CoxPH models confirmed the independent prognostic impact of CMSs in GBM patients after adjusting for other significant clinical/molecular factors in both GBM cohorts (TCGA-GBM: HR:1.457, 95% CI: 1.002-2.117, $P=0.049$; ZN-GBM: HR: 3.101, 95% CI: 1.681-5.72, $P<0.001$; Fig. 5e, f, and Supplementary Tables 16 and 17).

Lastly, we performed a pooled analysis by combining all LGG and GBM patients into Pooled-LGG (595 patients in total) and Pooled-GBM (459 patients in total) cohorts, respectively. The pooled analysis confirmed (1) the significantly worse OS of subtype 2 patients (Pooled-LGG: $P<0.0001$, Extended Data Fig. 10a; Pooled-GBM: $P<0.0001$, Extended Data Fig. 10c); and (2) the independent prognostic impact of CMSs in both pooled cohorts (Pooled-LGG: HR: 2.315, 95% CI: 1.617-3.315, $P<0.001$, Extended Data Fig. 10b and Supplementary Table 18; Pooled-GBM: HR: 1.521, 95% CI: 1.173-1.974, $P=0.002$, Extended Data Fig. 10d and Supplementary Table 19). Interestingly, OS difference between LGG subtypes was independent of tumor grade (Grade2: $P=0.037$; Grade3: $P<0.0001$; Extended Data Fig. 10e) and histology types (Astrocytoma: $P=0.0046$, Oligodendroglioma: $P=0.012$, Oligoastrocytoma: $P=0.0013$; Extended Data Fig. 10f), further demonstrating the independent clinical value of CMSs.

Molecular Annotation underlying Cellular Morphometric subtypes

To get insight into molecular differences underlying the CMSs, we used available transcriptome data from the TCGA-LGG and identified a total of 316 genes that are differentially expressed between CMSs ($|\log_2FC|>1$, $P<0.001$, Fig. 6a, and Supplementary Table 20), where 147 and 169 genes were upregulated and downregulated, respectively, in subtype 2 compared to subtype 1. Gene ontology (GO) functional enrichment analysis of the differentially expressed genes (DEGs) demonstrated significant enrichment ($FDR<0.05$) for biological processes involved hemostasis, keratinization, intermediate filament

organization, humoral immune response, regulation of ERK1 and ERK2 cascade, positive regulation of acute inflammatory response. (Fig. 6b, and Supplementary Table 21); Cellular component GO terms significantly enriched (FDR<0.05) in the DEGs included intermediate filament, blood microparticle, cluster of actin-based cell projections, collagen-containing extracellular matrix and trans-Golgi network transport vesicle (Fig. 6c, and Supplementary Table 22), whereas molecular function GO terms (FDR<0.05) included structural constituent of cytoskeleton and cytokine activity (Fig. 6d, and Supplementary Table 23). KEGG analysis indicated that DEGs were significantly enriched (FDR<0.05) in neuroactive ligand-receptor interaction, cytokine-cytokine receptor interaction, IL-17 signaling pathway, complement and coagulation cascades and Staphylococcus aureus infection (Fig. 6e, and Supplementary Table 24). Together these findings suggest possible differences in molecular mechanisms of two CMSs.

Association of Cellular Morphometric Subtypes with Tumor Immune Microenvironment

Based on the molecular annotation of DEGs between CMSs, we investigated their association with the immune microenvironment. Our study showed (Fig. 7) significantly more B cells (P=0.027), dendritic cells (P=0.024), eosinophils (P=0.033), macrophages (P=0.02), mast cells (P=0.0034), NK cells (P=0.01), neutrophils (P=0.025), gamma delta T cells (P=0.0097), T regulatory cells (P=0.0042), macrophages M1 (P=0.003), and monocytes (P=0.029) infiltrated in subtype 2. Meanwhile, there was a trend of an increased abundance of CD4⁺ T cells (P=0.065), CD8⁺ T cells (P=0.057), and plasma cells (P=0.072) in subtype 2. Moreover, the T cell infiltration score (P=0.00097) and overall immune infiltration score (P=0.029) were significantly higher in subtypes 2 (Fig. 7). To reveal the possibility of immune escape in subtype 2 patients, we examined the expression levels of two immune inhibitory receptor CTLA4 and PD-1 and the ligand of PD-1 (i.e., PD-L1). In the TCGA-LGG cohort (Fig. 8a), the expression of PD-1 (P=0.00044), PD-L1 (P=0.03), and CTLA4 (P=0.17) were significantly or tended to be higher in patients with subtype 2. Finally, we validated the expression levels of these immune inhibitory molecular markers in the ZN-LGG cohort using IHC, and the results confirmed the significant upregulation of PD-1 (P=8e-05), PD-L1 (P=0.018) and CTLA4 (P=0.00089) in subtype 2 (Fig. 8b,c). Overall, these results indicated possible mechanisms for immune escape or immune tolerance in subtype 2 tumors, which could explain the poor prognosis of subtype 2 patients.

Discussion

In this study, we extracted the cellular morphometric biomarkers from WSIs of LGG patients through unsupervised representation learning strategy, and subsequently defined two CMSs based on these biomarkers. The reproducibility and robustness of CMSs was demonstrated in two independent validation cohorts of LGG patients, and the transferability of CMSs was demonstrated in two independent validation cohorts of GBM patients. Moreover, the importance of CMSs lies in its independent prognostic significance after adjusting for other clinical and molecular factors, the association with treatment response, and the relation to underlying molecular and phenotypic alterations. Furthermore, we constructed a nomogram for predicting the survival rate of LGG patients at 3 and 5 years based on important clinical factors, molecular factors, and CMS.

Different from many end-to-end CNN-like systems, which mainly focuses on the prediction of clinical/molecular knowledge, the emphasis of present study focuses on novel knowledge discovery with interpretability, robustness and independent clinical value through multicentric validation. As a further justification, we evaluated a superior CNN-like system (i.e., SCNN: survival convolutional neural networks), which was specifically designed and optimized for the prediction of cancer outcomes from brain tumor histology and genomics²². Interestingly, SCNN risk score does not provide independent and significant prognostic value on both TCGA-LGG (P=0.182, Extended Data Fig. 9b) and TCGA-GBM (P=0.533, Extended Data Fig. 9c) cohorts, in the presence of our CMS subtype and other important clinical/molecular factors, which suggest that our patient stratification strategy **out-performs** the supervised CNN-like system (i.e., SCNN) for precision prognosis.

SCNA score, closely related to the occurrence and progression of many tumors (including glioma), is related to poor prognosis^{23,24}. Meanwhile, TMB levels, closely related to degree of malignancy and poor prognosis of glioma, is often used as a biomarker predicting the efficacy of anti-PD1 therapy²⁵⁻²⁷. And our study confirmed the significantly higher SCNA score and TMB level in subtype 2 patients, which explains the poor prognosis, and provides justification for further investigation of anti-PD-1 immunotherapy for subtype 2 patients, especially when subtype 2 patients were associated with poor response to primary therapy and follow-up treatment.

The tumor immune microenvironment is closely related to the progression of tumor cells. In the glioma microenvironment, NK cells, macrophages, neutrophils, CD4⁺ T cells, CD8⁺ T cells, regulatory T cells, etc. influence disease outcome²⁸. Our study showed, T cells (including CD4⁺ T cells, CD8⁺ T cells, gamma delta T Cell, regulatory T cells), B cells, plasma cells, macrophages, NK cells, neutrophils, mast cells, etc. were higher in subtype type 2 patients, suggesting a higher immune infiltration in subtype 2 patients. Given the significantly higher expression levels of PD-1, PD-L1, and CTLA4 as well as the relatively high TMB in subtype 2 patients, subtype 2 patients may benefit from anti-PD-1 immunotherapy.

CTLA4 inhibits T cell activation by inducing antigen-presenting cells to express CD80 and CD86²⁹. Regulatory T cells can inhibit T cell function by secreting IL-10 and TGF- β ³⁰. Studies have reported that neutrophils infiltration in tumor tissues can promote tumor progression and metastasis, and in glioma, neutrophils can promote tumor proliferation by inducing angiogenesis³¹⁻³³. NK cells are important component of the human immune system. However, Poli et al. showed that NK cells are in a state of inactivation in glioma³⁴. The above analysis of immune cell components showed that although subtype 2 is enriched for immune cells, due to some immunosuppressive cell (regulatory T cell) infiltration, T cell function inactivation and other factors, subtype 2 may be in immune tolerance.

This study has some shortcomings. First, relatively few LGG patients were included in the independent ZN-LGG and SU-LGG validation cohorts, so the conclusions of this study need to be further verified in large scale clinical studies. Second, our findings raise the possibility that subtype 2 LGG patients could benefit from anti-PD-1 immunotherapy; however, since LGG patients have not been recommended for anti-

PD-1 immunotherapy based on existing clinical practice, we could not find any retrospective dataset to test this and will investigate it in our future prospective study.

In conclusion, we developed a pathology-image-based LGG subtyping, that seems to stratify LGG patients into two groups with different OS associated with treatment responses, copy number alterations and TMB levels and immune tolerance.

Methods

Data collection

The patient data in this retrospective study, including diagnostic slides of tissue histology and the corresponding clinical information, were collected from TCGA-LGG cohort (Supplementary Table 1), Zhongnan Hospital of Wuhan University (ZN-LGG cohort, between January 2016 to May 2019, Supplementary Table 2), the Medical Center of Stanford University (SU-LGG cohort, between January 2012 to December 2013, Supplementary Table 3), TCGA-GBM cohort (Supplementary Table 4), and Zhongnan Hospital of Wuhan University (ZN-GBM cohort, between January 2016 to May 2019, Supplementary Table 5) to form the discovery cohort and the multi-center independent validation cohorts, respectively. The inclusion criteria were primary LGG and GBM with diagnostic slides and overall survival information available. This study was approved by the institutional review board (IRB) of both Zhongnan Hospital of Wuhan University and Stanford University.

Definition of treatment response in TCGA-LGG cohort

According to published literature³⁵, the treatment response in TCGA project was assessed using Response Evaluation Criteria in Solid Tumors (RECIST) measure that evaluates the size of measurable lesions of solid tumors³⁶. Based on RECIST, the patient's response is divided into four registrations: complete remission, partial remission, progressive disease and stable disease. Since we focus on whether the patients responded to treatments, the treatment response in TCGA-LGG cohort is categorized as: Response (including complete remission and partial remission), and non-Response (including progressive disease and stable disease) groups.

Identification of Cellular Morphometric Biomarkers

We developed an unsupervised feature learning pipeline based on SPSPD²⁰ for the unsupervised discovery of underlying cellular morphometric characteristics from the 15 cellular morphometric features extracted from the diagnostic slides of TCGA-LGG cohort. And thereafter identified 256 cellular morphometric biomarkers for cellular object representation. Specifically, in this study, we used single network layer with 256 dictionary elements (i.e., cellular morphometric biomarkers) and sparsity constraint 30 at a fixed random sampling rate of 1000 cellular objects per diagnostic slide from TCGA-LGG cohort (Extended Data Fig. 2a), where the network parameters (i.e., dictionary size and sparsity) were experimentally optimized to maintain the data reconstruction error ratio under certain threshold (i.e., 10%

in this study, Extended Data Fig. 2b,c). The pre-trained SPSP model reconstructs each cellular region (represented as a vector of 15 morphometric properties) as a sparse combination of pre-identified 256 cellular morphometric biomarkers, and thereafter represents each cellular object as the sparse code (i.e., sparse coefficients) during reconstruction, where the sparsity constraint leads to the reconstruction mainly contributed by the top 30 cellular morphometric biomarkers.

Clinical and Biological Evaluation of Cellular Morphometric Biomarkers

We evaluated the prognostic impact of top 30 out of 256 CMBs with largest variations mined from TCGA-LGG cohort with Cox proportional hazards regression (CoxPH) model (survival package in R, Version 3.2-3), and examined the effects of high or low levels of each prognostic significant CMB on OS using Kaplan-Meier analysis (survminer package in R, Version 0.4.8) and log-rank test (survival package in R, Version 3.2-3), where TCGA-LGG cohort was divided into CMB-high and CMB-low groups per CMB (survminer package in R, Version 0.4.8). Meanwhile, we evaluated biological significance between these CMB-high and CMB-low groups by assessing their relationship with biological factors available in the TCGA-LGG cohort using Mann-Whitney test.

Construction of Patient-Level Cellular Morphometric Context Representation

The patient-level cellular morphometric context representation was constructed based on pre-identified 256 cellular morphometric biomarkers as an aggregation (i.e., max-pooling) of all the cellular sparse codes extracted via pre-built SPSP model from the cellular objects belonging to the diagnostic slides of the same patients. Specifically, it consists of steps as follows, (1) delineation of cellular architecture and extraction of cellular morphometric properties from diagnostic slides of each patient; (2) construction of cellular sparse codes for the cellular objects belonging to each patient based on pre-identified 256 cellular morphometric biomarkers and pre-built SPSP model; (3) aggregation (i.e., max-pooling) of all cellular sparse codes belonging to the same patient to form the patient-level cellular morphometric representation; and (4) selection of the top 30 cellular morphometric biomarkers with the largest variations identified in TCGA-LGG cohort as the final patient-level cellular morphometric representation.

Identification and Application of Patient Subtype

The patient subtype was identified based on patient-level cellular morphometric context representation through consensus clustering strategy³⁷ (ConsensusClusterPlus package in R, Version 1.50.0) with hierarchical clustering, Pearson's correlation and 500 bootstrapping iterations; and the optimal number of subtypes was determined by the consistency of cluster assignment (consensus matrix) and the prognostic impact of subtypes. For a new patient, the subtype was assigned as follows: (1) construct patient-level cellular morphometric context representation with pre-built cellular morphometric biomarkers and SPSP model; (2) calculate the Pearson's distances between the new patient's representation and the mean representation of each pre-identified patient subtype; and (3) assign the new patient to its closest subtype yielding smallest Pearson's distance.

Clinical Evaluation and Validation of Patient Subtype

We evaluated and independently validated the clinical impact of pre-identified patient subtype from TCGA-LGG cohort, ZN-LGG cohort, SU-LGG cohort, TCGA-GBM cohort, and ZN-GBM cohort, respectively, where the latest clinical data of TCGA-LGG and TCGA-GBM cohorts was downloaded from Genomic Data Commons (GDC, <https://portal.gdc.cancer.gov/>), and the subtype assignment of each patient in independent validation cohorts (i.e., ZN-LGG, SU-LGG, TCGA-GBM, and ZN-GBM) was achieved through the application of pre-built TCGA-LGG patient subtype model as described previously. The evaluation and validation reside in three folds as follows, (1) Prognostic impact. The prognostic impact of patient subtype on OS was evaluated on TCGA-LGG, ZN-LGG, SU-LGG, TCGA-GBM, and ZN-GBM cohorts with univariate and stepwise multivariate Cox proportional hazards regression (CoxPH) models (survival package in R, Version 3.2-3), and the subtype-specific survival was visualized through Kaplan-Meier curve (survminer package in R, Version 0.4.8); (2) Predictive power of survival. A nomogram, based on multivariate CoxPH model, was developed to assist the prediction of 3-year and 5-year survival rate of LGG patients, where the multivariate CoxPH model was constructed with selected variables (i.e., clinical factors, molecular factors, and patient subtype) based on their significant and independent prognostic impact. Specifically, during nomogram construction and validation, the patients in TCGA-LGG cohort were randomly partitioned into training set (60% patients) and testing set (40% patients) through stratified sampling strategy. Then, a nomogram was constructed (rms package in R, Version 6.0-1) on the training set to predict the 3-year, and 5-year overall patient survival. The performance of nomogram was evaluated based on concordance-index (C-index) with 1000 bootstraps on TCGA-LGG training set and test set, followed by calibration analysis to calibrate the performance of the nomogram; and (3) Treatment response. The treatment response was categorized as: Response (including complete remission and partial remission); and Non-response (including progressive disease and stable disease). And the differences in treatment response were assessed with Chi-square test for both primary therapy and follow-up treatment.

Differences in Gene expression, Mutation load, and Immune microenvironment between Subtypes

Differentially expressed genes (DEGs) between patient subtypes were estimated (edgeR package in R, Version 3.30.3) based on the count data of TCGA-LGG cohort, where genes with $|\log_2FC| > 1$ (FC: fold change) and $P < 0.001$ were selected and visualized via volcano plot (EnhancedVolcano package in R, Version 1.6.0). Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis were performed³⁸ (clusterProfiler package in R, Version 3.16.1) to exam the biological functions of DEGs. The total mutation number and somatic copy number alteration (SCNA) of each TCGA-LGG sample were calculated (maftool package in R, Version 2.4.05)³⁹ on the basis of MuSe⁴⁰ preprocessed mutation data. The SCNA levels of each patient in the TCGA-LGG cohort were calculated according to previous work⁴¹. The infiltration scores of 18 immune cells and overall immune infiltration score were estimated via R package “ConsensusTME” (version: 0.0.1.9000)⁴², and total T cell infiltration score was calculated according to the method introduced by Senbabaoglu et al.⁴³.

Immunohistochemical (IHC) Staining

IHC staining was carried out on 4- μ m sections of formalin-fixed and paraffin-embedded tissues according to the standard protocol on the entire ZN-LGG cohort (70 patients in total). Briefly, sections were dewaxed and rehydrated in serial alcohol washes, and then the endogenous peroxidase activities were blocked. After the nonspecific sites were saturated with 5% normal goat serum, the sections were incubated overnight at 4°C with anti-PD-1 Ab (1:50, mouse mAb, #UMAB199, ZSGB-Bio), anti-PD-L1 Ab (1:100, rabbit mAb, #13684, Cell signaling), or anti-CTLA4 Ab (1:50, mouse mAb, #UMAB249, ZSGB-Bio), and then incubated with anti-rabbit or anti-mouse Ig secondary Ab. The sections were visualized with the biotin-peroxidase complex and were counterstained with hematoxylin. For the assessment of PD-1 and CTLA4, the stained sections were screened at low-power field ($\times 40$), and 5 hot spots were selected. The number of positive cells in these areas were counted at HPF $\times 400$, 0.47 mm². The expression of PD-L1 was scored as a percentage of tumor cells expressing PD-L1 (3, $\geq 50\%$; 2, $\geq 5\%$ and $< 50\%$; 1, $\geq 1\%$ and $< 5\%$; and 0, $< 1\%$), where the staining in areas of necrosis was not quantified. The assessment was conducted by two experienced neuropathologists blinded to clinical information.

Statistical Analysis

Survival differences between subtypes or groups were examined using log-rank test. Differences in the treatment response of primary therapy and follow-up treatment between subtypes were examined using Chi-square test. Differences in respect of the expression of four negative immune regulators CTLA4, PD-1 and PD-L1, the immune cell infiltration, and genomic heterogeneity (tumor mutation burden, somatic copy number alteration) between subtypes were analyzed with Mann-Whitney non-parametric test. P value (FDR corrected if applicable) less than 0.05 was considered to be statistically significant. All analysis was performed with R (Version 4.0.2).

Declarations

Acknowledgement: We thank Dr. Lee A. Cooper for providing SCNN Risk Score for our independent evaluation, as well as his thoughtful suggestions. This work was supported by National Cancer Institute (NCI) at the National Institutes of Health (NIH): R01CA184476 (H.C.). Additional support was provided by Zhongnan Hospital of Wuhan University: ZLYNXM202011 (Z.-Q.L.) for independent validation at Zhongnan Hospital. Lawrence Berkeley National Laboratory (LBNL) is a multi-program national laboratory operated by the University of California for the DOE under contract DE AC02-05CH11231.

Author Contributions

H.C. conceived and designed the overall study, acquired funding, led the multi-center research activity, and co-wrote the manuscript. H.V. and Z.-Q.L. led the independent validation in the corresponding hospital, supervised for the clinical data and WSIs acquisition in the corresponding validation cohort and co-wrote

the manuscript. X.L., J.-H. M., X.Y. and H.C. developed the algorithm, computational pipeline, and performed the computational analysis and statistical interpretation. X.J. and S.A. prepared the study cohorts for independent validation. S.-F.T. and Y.-X.C. performed IHC staining and assessment of gene expression. Y.X., X.J., S.A., A.M.S., P.J.D. and J.-H.M. interpreted results and co-wrote the manuscript. K.C. maintained the computational server and data. All authors read and approved the manuscript.

Competing interests: The authors declare no conflicts of interest

Data and materials availability: TCGA-LGG and TCGA-GBM data are available at TCGA-LGG and TCGA-GBM projects, respectively; Patient-level cellular morphometric context data as well as the corresponding clinical data are available as supplementary materials; WSIs from ZN-GBM, ZN-LGG and SU-LGG cohorts are available upon request to the corresponding authors after the IRB approval from the corresponding hospitals; Matlab code for cellular morphometric biomarker detection and application with pre-trained SPSPD model will be available at <https://bbds.lbl.gov/project/lgg-cms> upon the acceptance of manuscript.

References

1. Sidaway, P. Low-grade glioma subtypes revealed. *Nat Rev Clin Oncol* **17**, 335 (2020).
2. Pineros, M., Sierra, M.S., Izarzugaza, M.I. & Forman, D. Descriptive epidemiology of brain and central nervous system cancers in Central and South America. *Cancer Epidemiol* **44 Suppl 1**, S141-S149 (2016).
3. Sturm, D., Pfister, S.M. & Jones, D.T.W. Pediatric Gliomas: Current Concepts on Diagnosis, Biology, and Clinical Management. *J Clin Oncol* **35**, 2370-2377 (2017).
4. Louis, D.N., *et al.* The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathologica* **131**, 803-820 (2016).
5. Renzi, S., *et al.* Causes of death in pediatric neuro-oncology: the sickkids experience from 2000 to 2017. *J Neurooncol* **149**, 181-189 (2020).
6. Keshri, V., *et al.* Risk Stratification in Low Grade Glioma: A Single Institutional Experience. *Neurol India* **68**, 803-812 (2020).
7. Viacoz, A., Lekoubou, A. & Ducray, F. Chemotherapy in low-grade gliomas. *Curr Opin Oncol* **24**, 694-701 (2012).

8. Ceccarelli, M., *et al.* Molecular Profiling Reveals Biologically Discrete Subsets and Pathways of Progression in Diffuse Glioma. *Cell* **164**, 550-563 (2016).
9. Louis, D.N., *et al.* The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathol* **131**, 803-820 (2016).
10. Gesperger, J., *et al.* Improved Diagnostic Imaging of Brain Tumors by Multimodal Microscopy and Deep Learning. *Cancers (Basel)* **12**(2020).
11. Ozcan, H., *et al.* A comparative study for glioma classification using deep convolutional neural networks. *Math Biosci Eng* **18**, 1550-1572 (2021).
12. Mzoughi, H., *et al.* Deep Multi-Scale 3D Convolutional Neural Network (CNN) for MRI Gliomas Brain Tumor Classification. *J Digit Imaging* **33**, 903-915 (2020).
13. Ning, Z., *et al.* Multi-modal magnetic resonance imaging-based grading analysis for gliomas by integrating radiomics and deep features. *Ann Transl Med* **9**, 298 (2021).
14. Yang, Y., *et al.* Glioma Grading on Conventional MR Images: A Deep Learning Study With Transfer Learning. *Front Neurosci* **12**, 804 (2018).
15. Li, Z., Wang, Y., Yu, J., Guo, Y. & Cao, W. Deep Learning based Radiomics (DLR) and its usage in noninvasive IDH1 prediction for low grade glioma. *Sci Rep* **7**, 5467 (2017).
16. Fukuma, R., *et al.* Prediction of IDH and TERT promoter mutations in low-grade glioma from magnetic resonance images using a convolutional neural network. *Sci Rep* **9**, 20311 (2019).
17. Chang, H., *et al.* Invariant delineation of nuclear architecture in Glioblastoma Multiforme for clinical and molecular association. *IEEE Transactions on Medical Imaging* **32**, 670-682 (2013).
18. Chang, H., AD, B., Spellman, P. & Parvin, B. Classification of Tumor Histology via Morphometric Context. in *IEEE Computer Vision and Pattern Recognition* in press (2013).
19. Chang, H., *et al.* Morphometric analysis of TCGA glioblastoma multiforme. *BMC Bioinformatics* **12**, 484 (2011).
20. Chang, H., *et al.* Stacked Predictive Sparse Decomposition for Classification of Histology Sections. *Int J Comput Vis* **113**, 3-18 (2015).
21. Charles, N.A., Holland, E.C., Gilbertson, R., Glass, R. & Kettenmann, H. The brain tumor microenvironment. *Glia* **59**, 1169-1180 (2011).
22. Mobadersany, P., *et al.* Predicting cancer outcomes from histology and genomics using convolutional networks. *Proceedings of the National Academy of Sciences* **115**, E2970 (2018).

23. Zhang, L., *et al.* Genomic analysis of primary and recurrent gliomas reveals clinical outcome related molecular features. *Sci Rep* **9**, 16058 (2019).
24. Petralia, F., *et al.* Integrated Proteogenomic Characterization across Major Histological Types of Pediatric Brain Cancer. *Cell* **183**, 1962-1985 e1931 (2020).
25. Wang, L., *et al.* Tumor mutational burden is associated with poor outcomes in diffuse glioma. *BMC Cancer* **20**, 213 (2020).
26. Marabelle, A., *et al.* Association of tumour mutational burden with outcomes in patients with advanced solid tumours treated with pembrolizumab: prospective biomarker analysis of the multicohort, open-label, phase 2 KEYNOTE-158 study. *Lancet Oncol* **21**, 1353-1365 (2020).
27. Johnson, A., *et al.* Comprehensive Genomic Profiling of 282 Pediatric Low- and High-Grade Gliomas Reveals Genomic Drivers, Tumor Mutational Burden, and Hypermutation Signatures. *Oncologist* **22**, 1478-1490 (2017).
28. Qi, Y., Liu, B., Sun, Q., Xiong, X. & Chen, Q. Immune Checkpoint Targeted Therapy in Glioma: Status and Hopes. *Front Immunol* **11**, 578877 (2020).
29. Fong, B., *et al.* Monitoring of regulatory T cell frequencies and expression of CTLA-4 on T cells, before and after DC vaccination, can predict survival in GBM patients. *PLoS One* **7**, e32614 (2012).
30. Bettelli, E., *et al.* Reciprocal developmental pathways for the generation of pathogenic effector TH17 and regulatory T cells. *Nature* **441**, 235-238 (2006).
31. Khan, S., *et al.* Role of Neutrophils and Myeloid-Derived Suppressor Cells in Glioma Progression and Treatment Resistance. *Int J Mol Sci* **21**(2020).
32. Garrido-Navas, C., *et al.* Cooperative and Escaping Mechanisms between Circulating Tumor Cells and Blood Constituents. *Cells* **8**(2019).
33. Fujita, M., *et al.* Role of type 1 IFNs in antiglioma immunosurveillance—using mouse studies to guide examination of novel prognostic markers in humans. *Clin Cancer Res* **16**, 3409-3419 (2010).
34. Poli, A., *et al.* Targeting glioblastoma with NK cells and mAb against NG2/CSPG4 prolongs animal survival. *Oncotarget* **4**, 1527-1546 (2013).
35. Piñeiro-Yáñez, E., Jiménez-Santos, M.J., Gómez-López, G. & Al-Shahrour, F. In Silico Drug Prescription for Targeting Cancer Patient Heterogeneity and Prediction of Clinical Outcome. *Cancers (Basel)* **11**(2019).
36. Wolchok, J.D., *et al.* Guidelines for the evaluation of immune therapy activity in solid tumors: immune-related response criteria. *Clin Cancer Res* **15**, 7412-7420 (2009).

37. McKenna, S.J., Ricketts, I.W., Cairns, A.Y. & Hussein, K.A. A comparison of neural network architectures for cervical cell classification. in *1993 Third International Conference on Artificial Neural Networks* 105-109 (1993).
38. Yu, G., Wang, L.G., Han, Y. & He, Q.Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284-287 (2012).
39. Mayakonda, A., Lin, D.C., Assenov, Y., Plass, C. & Koeffler, H.P. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res* **28**, 1747-1756 (2018).
40. Fan, Y., *et al.* MuSE: accounting for tumor heterogeneity using a sample-specific error model improves sensitivity and specificity in mutation calling from sequencing data. *Genome Biol* **17**, 178 (2016).
41. Davoli, T., Uno, H., Wooten, E.C. & Elledge, S.J. Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. *Science* **355**(2017).
42. Jiménez-Sánchez, A., Cast, O. & Miller, M.L. Comprehensive Benchmarking and Integration of Tumor Microenvironment Cell Estimation Methods. *Cancer Res* **79**, 6238-6246 (2019).
43. Şenbabaoğlu, Y., *et al.* Tumor immune microenvironment characterization in clear cell renal cell carcinoma identifies prognostic and immunotherapeutically relevant messenger RNA signatures. *Genome Biol* **17**, 231 (2016).

Figures

Figure 1

Graphical illustration of our study design with knowledge mining and double-blind validation.

Figure 2

Unsupervised feature learning discovers cellular morphometric biomarkers with clinical significance and molecular relevance. (a) Prognostic significant CMBs with favorable and unfavorable examples and LGG patients within CMB-low and CMB-high groups (cutoffs summarized in Supplementary Table 12; (b) significant difference in various tumor microenvironmental factors; (c) significant difference in survival. (d) significant difference in genetic instability.

Figure 3

Figure 3

LGG patient subtype provides significant and independent prognostic impact. (a) Consensus clustering model for LGG patient subtypes discovery and inference; (b) Subtype specific patients in TCGA-LGG cohort form distinct clusters in patient level cellular morphometric context space; (c) Subtype specific patients in TCGA-LGG cohort show significant difference in survival; (d) Patient subtype in TCGA-LGG cohort is a significant and independent prognostic factor; (e) Subtype specific patients in ZN-LGG cohort

form distinct clusters in patient level cellular morphometric context space; (f) Subtype specific patients in ZN-LGG cohort show significant difference in survival; (g) Patient subtype in ZN-LGG cohort is a significant and independent prognostic factor; (h) Subtype specific patients in SU-LGG cohort form distinct clusters in patient level cellular morphometric context Space; (i) Subtype specific patients in SU-LGG cohort show significant difference in survival; (j) Patient subtype in SU-LGG cohort is a significant and independent prognostic factor.

Figure 4

Figure 4

Development and validation of nomogram predicting the 3- year and 5-year survival of LGG patients. (a) Nomogram predicting the 3- year and 5-year survival of LGG patients. (b) Calibration analysis at 3- year in the training set of TCGA-LGG cohort. (c) Calibration analysis at 5- year in the training set of TCGA-LGG cohort. (d) Calibration analysis at 3- year in the test set of TCGA-LGG cohort. (e) Calibration analysis at 5- year in the test set of TCGA-LGG cohort.

Figure 5

Figure 5

Application of LGG cellular morphometric biomarkers and subtypes to GBM cohorts leads to distinct patient stratification (a-b) with significant differences in OS (c-d) and independent prognostic values on TCGA-GBM (e) and ZN-GBM (f) cohorts, respectively. Note, G-CIMP and MGMT status are non-significant prognostic factors via univariate CoxPH analysis in ZN-GBM cohort, and therefore was not included in multivariate CoxPH analysis as shown in (f).

Figure 6

Differentially expressed genes (DEG) between two subtype and associated functional enrichment analyses. (a) Volcano plot depicting the differentially expressed genes with $|\log_2 FC| > 1$ and $P < 0.001$. (b) Biological process (BP) enrichment analysis on DEGs. (c) Cellular component (CC) enrichment analysis on DEGs. (d) Molecular function (MF) enrichment analysis on DEGs. (e) Kyoto Encyclopedia of Genes and Genomes pathway enrichment analysis on DEGs.

Figure 7

Figure 7

Patient subtypes in TCGA-LGG cohort show significant difference in various tumor microenvironmental factors.

Figure 8

Immunohistochemical (IHC) staining confirms the upregulation of PD-1, PD-L1 and CTLA 4 in subtype 2 LGG patients. (a) Subtype specific expression of PD-1 (first row), PD-L1 (second row) and CTLA4 (third row) in TCGA-LGG cohort. (b) Representative examples of PD-1 staining (first row), PD-L1 staining (second row) and CTLA4 staining (third row) in Subtype 1 and 2 LGG patients, respectively, where PD-1 expression was frequently observed in the plasma of lymphocytes around blood vessels; PD-L1 was widely expressed in the membrane of tumor cells, while slightly in cytoplasm; and CTLA4 positive expression was majorly observed in the cytoplasm of lymphocytes around blood vessels. (c) Subtype specific expression of PD-1 (first row), PD-L1 (second row) and CTLA4 (third row) quantified via IHC staining in ZN-LGG cohort.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [TableS13.docx](#)
- [LGGmanuscriptExtendedDataFigures20210730.pdf](#)

- [TableS5.docx](#)
- [TableS6.docx](#)
- [TableS8.xlsx](#)
- [TableS17.docx](#)
- [TableS18.docx](#)
- [TableS12.xlsx](#)
- [TableS10.xlsx](#)
- [TableS15.docx](#)
- [TableS7.xlsx](#)
- [TableS19.docx](#)
- [TableS11.xlsx](#)
- [TableS4.docx](#)
- [TableS2.docx](#)
- [TableS1.docx](#)
- [TableS3.docx](#)
- [TableS22.xlsx](#)
- [TableS21.xlsx](#)
- [TableS20.xlsx](#)
- [TableS14.docx](#)
- [TableS16.docx](#)
- [TableS24.xlsx](#)
- [TableS9.xlsx](#)
- [TableS23.xlsx](#)