

Original Paper

# Investigating Associations Between Prognostic Factors in Gliomas: Unsupervised Multiple Correspondence Analysis

Maria Eduarda Goes Job<sup>1</sup>; Heidge Fukumasu<sup>1</sup>, PhD; Tathiane Maistro Malta<sup>2</sup>, PhD; Pedro Luiz Porfirio Xavier<sup>1</sup>, PhD

<sup>1</sup>Laboratory of Comparative and Translational Oncology, Department of Veterinary Medicine, School of Animal Science and Food Engineering, University of Sao Paulo, Pirassununga, Brazil

<sup>2</sup>Cancer Epigenomics Laboratory, Department of Clinical Analysis, Toxicology and Food Sciences, School of Pharmaceutical Sciences of Ribeirao Preto, University of Sao Paulo, Ribeirão Preto, Brazil

**Corresponding Author:**

Pedro Luiz Porfirio Xavier, PhD

Laboratory of Comparative and Translational Oncology, Department of Veterinary Medicine

School of Animal Science and Food Engineering, University of Sao Paulo

Avenida Duque de Caxias, 225, Jardim Elite

Pirassununga, 13.635-900

Brazil

Phone: 55 19 9 8367-1821

Email: [porfirioxavier@usp.br](mailto:porfirioxavier@usp.br)

## Abstract

**Background:** Multiple correspondence analysis (MCA) is an unsupervised data science methodology that aims to identify and represent associations between categorical variables. Gliomas are an aggressive type of cancer characterized by diverse molecular and clinical features that serve as key prognostic factors. Thus, advanced computational approaches are essential to enhance the analysis and interpretation of the associations between clinical and molecular features in gliomas.

**Objective:** This study aims to apply MCA to identify associations between glioma prognostic factors and also explore their associations with stemness phenotype.

**Methods:** Clinical and molecular data from 448 patients with brain tumors were obtained from the Cancer Genome Atlas. The DNA methylation stemness index, derived from DNA methylation patterns, was built using a one-class logistic regression. Associations between variables were evaluated using the  $\chi^2$  test with k degrees of freedom, followed by analysis of the adjusted standardized residuals (ASRs >1.96 indicate a significant association between variables). MCA was used to uncover associations between glioma prognostic factors and stemness.

**Results:** Our analysis revealed significant associations among molecular and clinical characteristics in gliomas. Additionally, we demonstrated the capability of MCA to identify associations between stemness and these prognostic factors. Our results exhibited a strong association between higher DNA methylation stemness index and features related to poorer prognosis such as glioblastoma cancer type (ASR: 8.507), grade 4 (ASR: 8.507), isocitrate dehydrogenase wild type (ASR:15.904), unmethylated MGMT (methylguanine methyltransferase) Promoter (ASR: 9.983), and telomerase reverse transcriptase expression (ASR: 3.351), demonstrating the utility of MCA as an analytical tool for elucidating potential prognostic factors.

**Conclusions:** MCA is a valuable tool for understanding the complex interdependence of prognostic markers in gliomas. MCA facilitates the exploration of large-scale datasets and enhances the identification of significant associations.

*JMIR Bioinform Biotech* 2025;6:e65645; doi: [10.2196/65645](https://doi.org/10.2196/65645)

**Keywords:** brain tumors; bioinformatics; stemness; multiple correspondence analysis

## Introduction

Cancer is a dynamic and heterogeneous disease characterized by several hallmarks controlling and contributing to its development and progression [1]. Cancer research

continually generates large scales of data encompassing clinical information, genomic and transcriptomic profiles, prognostic and diagnostic markers, and therapeutic targets [2]. Different approaches have been used to study and associate all these variables to manage this complexity,

aiming to reduce the dimensionality and enhance data interpretation and decision-making process. Several features used to study and classify the different types of cancer are based on categorical variables. For instance, the most widely used cancer staging system, TNM, is based on categorical variables, where “T” refers to the size of the primary tumor, “N” refers to the number of lymph nodes affected by cancer, and “M” refers to absence or presence of metastasis [3]. Thus, these biological and clinical variables interact, and their associations can be measured and diagnosed using statistical tests such as Fisher exact tests and  $\chi^2$  tests. However, these approaches could not provide a global and comprehensive picture of the associations between these variables, particularly in datasets with a large number of categorical variables. Therefore, using multivariate and visual analysis methods can significantly improve the analysis and interpretation of associations between clinical and molecular cancer phenotypes.

Brain tumors are a particularly aggressive type of cancer, mostly due to local tissue damage and highly invasive growth. Gliomas, which originate from neuroglial stem cells or progenitor cells, account for 30% of primary brain tumors and 80% of malignant brain tumors [4]. This heterogeneous disease is histologically classified based on anaplasia criteria and predominant cell types such as oligodendroglioma, astrocytoma, and glioblastoma (GBM) [5]. Nevertheless, as further investigation aimed to elucidate the neuropathological mechanisms of gliomas, new variables are considered for characterizing this cancer tumor, leading to reclassifications based on mutational profiles, clinical data, and epigenetic factors [6]. This scenario resulted in different prognosis predictions, diagnosis determination, and treatment responses, contributing to an increasingly complex and stratified understanding of gliomas.

Stemness is a key phenotype of cancer stem cells (CSCs), related to tumor initiation and progression, therapy resistance, and metastasis [7]. CSCs are referred to as a subpopulation of tumor cells able to self-renew and differentiate into distinct cell lineages, enabling those cells to adapt to different environmental situations [8]. Moreover, recent studies have demonstrated associations between stemness features and different histologic classifications or prognostic factors of gliomas [9-11]. Therefore, providing a comprehensive visualization of the associations between clinical features and stemness in brain tumors could be valuable for identifying and determining potential prognostic and therapeutic markers.

Multiple correspondence analysis (MCA) is an unsupervised data science methodology that aims to observe and represent associations between variables disposed in contingency tables, visualizing these associations in a 2D perceptual map. This approach allows for the simultaneous visualization of the relationship between 2 or more characteristics [12]. MCA shares general characteristics, and it is an extension of principal component analysis which is effective in reducing data dimensionality. Thus, MCA can significantly reduce the workload and simplify statistical analysis in healthy research [13]. The results of MCA are typically interpreted in a 2D map, where the relative positions

of categories of each variable and their distribution along the dimensions are analyzed. Categories that cluster together and are closer are more likely to be associated, providing key insights into the relationship [14]. Despite its applicability, rigor, and success in other disciplines such as Geography, Epidemiology, and Human Physiology, MCA remains underused in Oncology research and few studies are applying [12,14-16].

By using MCA, we aimed to gain a deeper understanding of the interdependence between stemness and prognostic factors. Our findings revealed associations among molecular and clinical characteristics and prognostic factors, as previously described by the literature [17]. Additionally, we demonstrated the capability of MCA to identify associations between stemness and these prognostic factors. Our results exhibited a strong association between higher stemness index and features related to poorer prognosis, demonstrating the utility of MCA as an analytical tool for elucidating oncological heterogeneity and may also offer a valuable strategy for therapeutic decision-making. This study highlights MCA as a powerful tool for overcoming the barrier of representing the heterogeneity and complexity of cancer variables, particularly in glioma.

## Methods

### *Dataset of the Tumor Samples*

Clinical and molecular information of a total of 448 patients with brain tumors was obtained from the Cancer Genome Atlas (TCGA). We tailored the dataset to contain only qualitative information, with 12 variables: cancer type, histology, grade, patient’s vital status, IDH (isocitrate dehydrogenase) status, codeletion of chromosomes 1p and 19q arms, MGMT (methylguanine methyltransferase) gene methylation, telomerase reverse transcriptase (TERT) expression, gain of chromosome 19 and 20, chromosome 7 gain and chromosome 10 loss, ATRX (alpha thalassemia/mental retardation syndrome, X-linked) status, and GBM transcriptome subtypes. All categorical variables were selected based on their established role as prognostic factors for brain tumors.

### *DNA Methylation Stemness Index*

The DNA methylation stemness index (mDNAsi) based on DNA methylation was built using a one-class logistic regression [18] on the pluripotent stem cell samples (embryonic stem cell and induced pluripotent stem cell) from the Progenitor Cell Biology Consortium dataset [19,20]. The algorithm was built and validated as described in the original paper [21]. The mDNAsi was applied in 381 samples from the TCGA database. Malta’s model presented a high correlation among other CSC signatures, providing significant insights into the biological and clinical features of pan-cancer. The workflow to generate the mDNAsi is available in the original paper [21].

## Multiple Correspondence Analysis

MCAs were conducted in the RStudio (version 4.3.1; Posit, PBC) environment using the packages FactoMineR (version 2.11; Institut Agro) [22] and cabooters (version 2.1.0; Cranfield University), for creating matrices for MCAs. Contingency tables for the categorical variables were generated, and associations between variables were assessed using a  $\chi^2$  test with  $k$  degrees of freedom. This was followed by the analysis of the adjusted standardized residuals (ASRs). The  $\chi^2$  test evaluates whether the observed associations between categorical variables are nonrandomly associated ( $P$  value  $<.05$ ). ASRs higher than 1.96 indicate a significant association between variables in the matrix. To perform MCA, the categorical variables should not be randomly associated. To create the perceptual map, inertia was determined as the total  $\chi^2$  divided by the number of samples, resulting in the number of associations in the dataset. MCA was performed based on the binary matrices and row and column profiles were determined to demonstrate the influence of each category of variables on the others. Matrices were defined based on the row and column profiles. Eigenvalues were then extracted to represent the number of dimensions that could be captured in the analysis. Finally, the x- and y-axis coordinates of the perceptual map were determined, allowing the category of the variables to be represented and established. In MCA, the spatial distance between categories of different variables reflects their associations. Categories with high coordinates that are close in space are directly associated, while categories presenting high coordinates but opposing coordinates are inversely associated.

## Statistical Analysis

Fisher exact tests and  $\chi^2$  tests were performed using RStudio 4.3.1 environment and GraphPad Prism (version 10.3.0; Dotmatics, USA).

## Ethical Considerations

The results published in this paper are in whole based upon data generated by the TCGA Research Network [23]. TCGA Ethics and Policies was originally published by the National Cancer Institute [24].

## Results

### MCA Can Identify Associations Between Different Variables of Gliomas and Patient Vital Status

To determine the suitability of glioma variables for MCA, we first evaluated whether categorical glioma variables

were randomly or nonrandomly associated. This involved creating individual contingency tables for each pair of glioma variables (Multimedia Appendices 1-13). Then, we applied  $\chi^2$  tests for each contingency table to confirm nonrandom associations ( $P$  value  $<.05$ ). We also confirmed the associations between categorical variables and patients' vital status using the Fisher exact test ( $P$  value  $<.05$ ) (Multimedia Appendix 14). Based on the  $\chi^2$  test, the results indicated that only 2 categorical variables, gender and DAXX expression, were randomly associated, suggesting no significant association patterns between these variables and the others. Consequently, gender and DAXX expression were excluded from further analysis.

In the subsequent analysis, we observed and measured the strength of associations between the patient vital status (0-alive; 1-dead) and different factors including cancer type, histology, grade, IDH status, 1p19q codeletion, MGMT promoter methylation, gain of chromosome (Chr) 7 and loss of Chr10 (7+/10-), co-gain of Chr19 and Chr20 (19+/20+), TERT expression, ATRX status, and transcriptome subtype, aiming to determine whether MCA could identify associations between prognostic factors for this disease. We used ASRs to assess these associations, considering a category of each variable to be associated with either alive or dead vital status when the ASR values were higher than 1.96. Patients' vital status classified as dead were associated with poorer prognostic factors such as GBMs, grade 4, IDH wild type, non-codeleted 1p19q, unmethylated MGMT promoter, gain of Chr7 and loss of Chr10, expression of TERT, ATRX wild type, and classical (CL) and mesenchymal (ME) transcriptome subtypes (Table 1). In contrast, patients classified as alive were linked to favorable prognostic variables, including oligoastrocytomas and oligodendrogliomas, grade 2, IDH mutant, codeleted 1p19q, methylated MGMT promoter, absence of combined Chr7+/Chr10- (chromosome 7 gain and 10 loss), lack of TERT expression, ATRX mutant, and the proneural (PN) and neural (NE) transcriptome subtypes (Table 1). Histological classification, grade, IDH status, and Chr7+/Chr10- were the most strongly associated features with patient vital status. These associations were further illustrated in a heatmap (Figure 1A-D).

**Table 1.** Table exhibiting the values of the adjusted standardized residuals. Categories of variables with values higher than 1.96 are considered associated. We could observe a strong association between poorer prognostic factors and dead vital status. In contrast, better prognostic factors were associated with alive vital status.

Glioma variables	Patient vital status		Categories associated with
	Alive	Dead	
Glioblastoma	— <sup>a</sup>	8.127	Dead
Oligoastrocytoma	2.64	—	Alive
Oligodendroglioma	3.309	—	Alive
Astrocytoma	1.756	—	Not associated
Grade 2	6.809	—	Alive
Grade 3	0.155	—	Not associated
Grade 4	—	8.127	Dead
IDH <sup>b</sup> wild type	—	8.804	Dead
IDH mutant	8.804	—	Alive
1p/19q codeletion	5.265	—	Alive
1p/19q non-codeletion	—	5.265	Dead
Methylated MGMT <sup>c</sup> promoter	5.26	—	Alive
Unmethylated MGMT promoter	—	5.26	Dead
No combined Chr7+/Chr10 <sup>-d</sup>	5.756	—	Alive
Chr7+/Chr10 <sup>-</sup>	—	5.756	Dead
Not expressed TERT <sup>e</sup>	3.078	—	Alive
Expressed TERT	—	3.078	Dead
ATRX <sup>f</sup> mutant	2.311	—	Alive
ATRX wild type	—	2.311	Dead
Proneural subtype	4.122	—	Alive
Neural subtype	3.593	—	Alive
Mesenchymal subtype	—	4.635	Dead
Classical subtype	—	4.852	Dead

<sup>a</sup>Not applicable.

<sup>b</sup>IDH: isocitrate dehydrogenase.

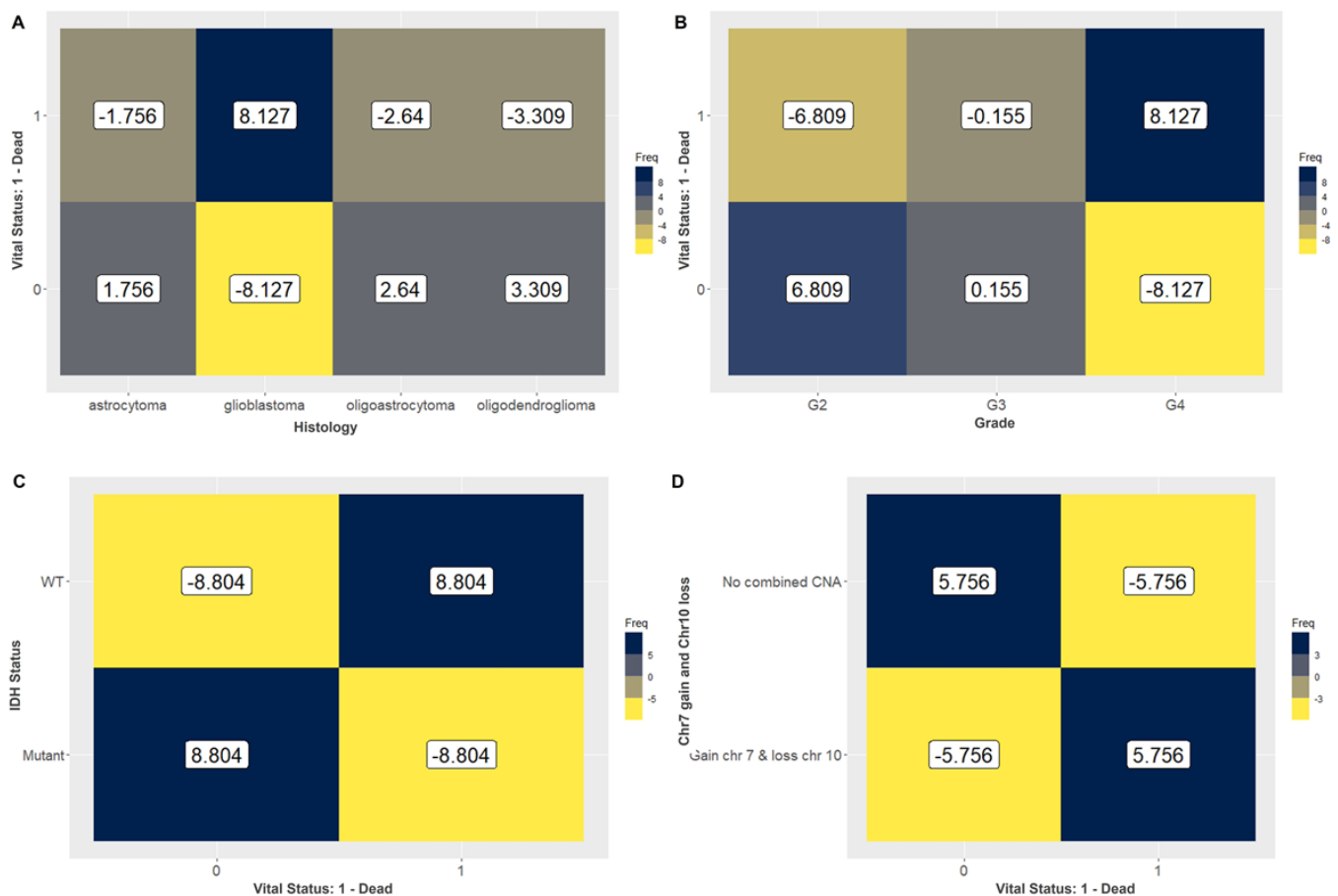
<sup>c</sup>MGMT: methylguanine methyltransferase.

<sup>d</sup>Chr7+/Chr10<sup>-</sup>: chromosome 7 gain and 10 loss.

<sup>e</sup>TERT: telomerase reverse transcriptase.

<sup>f</sup>ATRX: alpha thalassemia/mental retardation syndrome, X-linked.

**Figure 1.** Heatmap exhibiting the values of the adjusted standardized residuals. Categories of variables with values higher than 1.96 are associated with dead vital status. We could observe a strong association of (A) glioblastoma (8.127), (B) grade 4 (8.127), (C) IDH wild type (8.804), and (D) Chr7+/Chr10- (5.756) with dead vital status. Favorable prognostic factors including (A) oligoastrocytoma and oligodendroglioma, (B) grade 2, (C) IDH mutant, and (D) no combined copy number alterations were associated with alive vital status. Chr7+/Chr10-: chromosome 7 gain and 10 loss; IDH: isocitrate dehydrogenase.



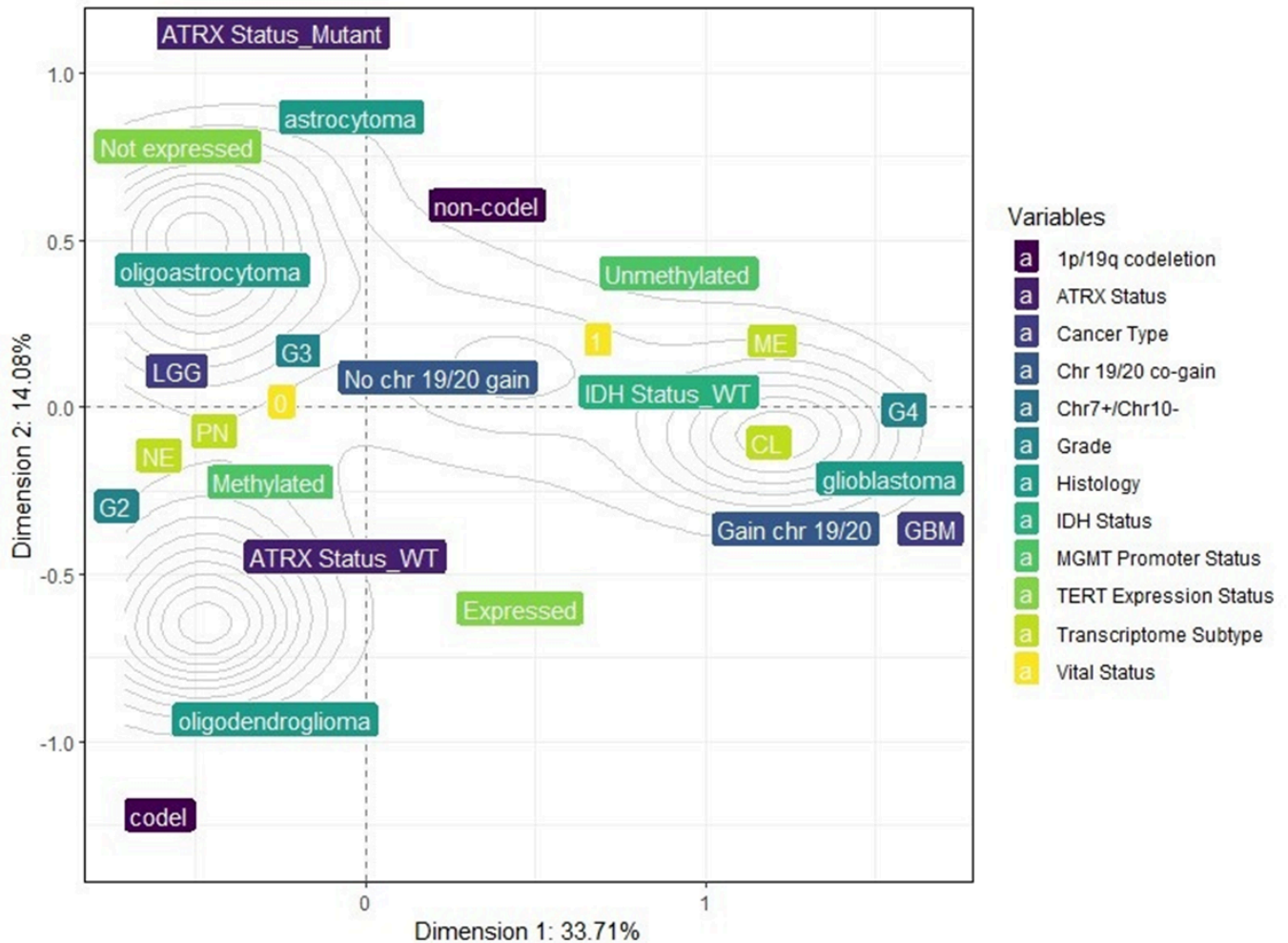
Using MCA, we observed that dimension 1 (x-axis) accounted for 33.71% of the variance, while dimension 2 (y-axis) accounted for 14.08%. The inertia (sum of the variances) for these 2 dimensions was 47.79%. The variance of the overall dimensions (17 dimensions) for the combinations of the variables is illustrated in [Multimedia Appendix 15](#). The main idea was to present the percentage of explained variance for each dimension and not the influence of individual variables. The total inertia (sum of the variances) was 1.41.

The results obtained from the MCA were visualized in a 2D perceptual map ([Figure 2](#)), highlighting the associations between the categories of each variable. The coordinates of each category are detailed in [Table 2](#). The perceptual map reveals that categories such as GBM, unmethylated MGMT promoter, IDH wild type, Chr7 gain and Chr10 loss, grade 4, GBM ATRX wild type, TERT expression, non-codel 1p.19q,

and CL and ME transcriptome subtypes are closely associated with dead vital status, appearing along the positive x-axis (dimension 1). Conversely, categories like oligoastrocytomas and oligodendrogliomas, grade 2, IDH mutant, codel 1p19q, methylated MGMT promoter, no combined copy number alterations, no expression of TERT, ATRX mutant, and PN and NE transcriptome subtypes are closely associated with alive vital status, appearing along the negative x-axis (dimension 1) ([Figure 2](#)).

These findings highlight the utility and capacity of MCA in reducing data dimensionality and demonstrate that, in gliomas, variables interact cohesively. MCA allows us to further visualize these interactions on a global perceptual map, organizing the characteristics into distinct clusters that correspond to different prognostic profiles.

**Figure 2.** Multiple correspondence analysis (MCA) 2D perceptual map demonstrating the association between the categories of each categorical variable. Categories that are closely clustered are strongly associated with each other. Categories such as glioblastoma, unmethylated MGMT promoter, IDH wild type, chromosome 7 gain and 10 loss (Chr7+/Chr10-), grade 4, glioblastoma ATRX wild type, TERT expression, non-codel 1p.19q, CL and ME transcriptome subtypes are closely associated with dead vital status (1), appearing along the positive x-axis (dimension 1). ATRX: alpha thalassemia/mental retardation syndrome, X-linked; CL: classical; GBM: glioblastoma; IDH: isocitrate dehydrogenase; ME: mesenchymal; MGMT: methylguanine methyltransferase; NE: neural; PN: proneural; TERT: telomerase reverse transcriptase.



**Table 2.** Coordinates of each category compounding the perceptual map.

Category	Dimension 1 (x-axis)	Dimension 2 (y-axis)
GBM <sup>a</sup>	1.6650830	-0.0896760
Low-grade glioma	-0.4723301	0.0254382
Astrocytoma	-0.2672355	0.9527631
Glioblastoma	1.6650830	-0.0896760
Oligoastrocytoma	-0.5334711	0.3276318
Oligodendroglioma	-0.6011671	-0.9346433
Grade 2	-0.6611308	-0.1971919
Grade 3	-0.2970898	0.2320783
Grade 4	1.6650830	-0.0896760
0-Alive	-0.3185609	-0.0551369
1-Dead	0.7544862	0.1305874
IDH <sup>b</sup> mutant	-0.6734117	-0.0548104
IDH wild type	1.1888626	0.0967641
1p/19q codel	-0.6877365	-13.034.766
1p/19q non-codel	0.2750946	0.5213906

Category	Dimension 1 (x-axis)	Dimension 2 (y-axis)
Methylated	-0.3429710	-0.1087842
Unmethylated	1.0048449	0.3187185
Chr7+/Chr10- <sup>c</sup>	1.4087248	-0.0210234
No combined Chr7+/Chr10-	-0.4205758	0.0062766
Chr 19/20 co-gain	1.4900007	-0.1295089
No Chr 19/20 co-gain	-0.0843397	0.0073307
Expressed TERT <sup>d</sup>	0.3715020	-0.6845760
Not expressed TERT	-0.4690682	0.8643636
ATRX <sup>e</sup> mutant	-0.6448249	1.0773395
ATRX wild type	0.2693572	-0.4500279
Classical	1.2675815	-0.0217510
Mesenchymal	1.0920361	0.2687642
Neural	-0.5475482	-0.0650952
Proneural	-0.5971662	-0.0604168

<sup>a</sup>GBM: glioblastoma.

<sup>b</sup>IDH: isocitrate dehydrogenase.

<sup>c</sup>Chr7+/Chr10-: chromosome 7 gain and 10 loss.

<sup>d</sup>TERT: telomerase reverse transcriptase.

<sup>e</sup>ATRX: Alpha Thalassemia/Mental Retardation Syndrome X-linked.

## MCA Can Associate an Epigenetic Stemness Index (mDNAsi) as a Prognostic Factor in Gliomas

After demonstrating that MCA effectively reduces dimensionality and identifies associations between prognostic factors and clinical data in the glioma database, we proceeded to explore whether MCA could also associate these variables with stemness phenotype. For this analysis, we updated our database by including mDNAsi as a new variable, categorized into low, intermediate, and high levels of stemness. These categories were based on the DNA methylation index related to tumor pathology and clinical outcomes, as previously studied by [21].

First, we evaluated whether the categorical glioma variables were randomly or nonrandomly associated with mDNAsi by creating individual contingency tables for each pair of glioma variables and applying  $\chi^2$  tests (Multimedia Appendix 16). We also confirmed the associations between

categorical variables using the Fisher exact test ( $P$  value  $<.05$ ) ( Multimedia Appendix 17). All the variables were found to be suitable for MCA. Then, using ASR values to evaluate the strength of these associations, our results indicated strong associations between high mDNAsi levels and poor prognostic and clinical factors. Higher mDNAsi levels were associated with GBM, IDH wild-type, absence of 1p19q co-deletion, unmethylated MGMT promoter, TERT expression, grade 3 and 4, patient's vital status as dead, Chr7+/Chr10-, chromosomes 19/20 co-gain, ATRX wildtype and ME and CL transcriptome subtypes (Table 3). Conversely, intermediate and lower levels of mDNAsi were associated with characteristics related to favorable prognosis, including oligodendroglioma, IDH mutant, 1p19q co-deletion, methylation of MGMT promoter, absence of TERT expression, grade 2, patient's vital status as alive, no combined copy number alteration, absence of chromosomes 19/20 co-gain, ATRX mutant, and PN and NE transcriptome subtypes (Table 3).

**Table 3.** Table exhibiting the values of the adjusted standardized residuals. Categories of variables with values higher than 1.96 are considered associated. We could observe a strong association between poorer prognostic factors and a higher stemness index (DNA methylation stemness index [mDNAsi]). In contrast, better prognostic factors were associated with lower stemness index.

Glioma Variables	mDNAsi			Categories associated with
	Low	Intermediate	High	
Glioblastoma	— <sup>a</sup>	—	8.507	High
Oligoastrocytoma	—	—	—	Not associated
Oligodendroglioma	3.949	—	—	Low
Astrocytoma	—	—	2.832	High
G2	3.279	4.057	—	Low and intermediate
G3	—	—	2.392	High
G4	—	—	8.507	High

Glioma Variables	mDNAsi			Categories associated with
	Low	Intermediate	High	
IDH <sup>b</sup> wild type	—	—	15.904	High
IDH mutant	8.743	7.057	—	Low and intermediate
1p/19q codeletion	5.772	2.102	—	Low and intermediate
1p/19q non-codeletion	—	—	7.964	High
Methylated MGMT <sup>c</sup> promoter	5.944	3.961	—	Low and intermediate
Unmethylated MGMT promoter	—	—	9.983	High
No combined Chr7+/Chr10 <sup>-d</sup>	6.436	5.927	—	Low and intermediate
Chr7+/Chr10 <sup>-</sup>	—	—	12.433	High
Not expressed TERT <sup>e</sup>	—	3.216	—	Intermediate
Expressed TERT	—	—	3.351	High
ATRX <sup>f</sup> mutant	—	3.505	—	Intermediate
ATRX wild type	—	—	4.949	High
Proneural subtype	8.476	—	—	Low
Neural subtype	—	4.218	—	Intermediate
Mesenchymal subtype	—	—	4.771	High
Classical subtype	—	—	10.981	High

<sup>a</sup>Not applicable.

<sup>b</sup>IDH: isocitrate dehydrogenase.

<sup>c</sup>MGMT: methylguanine methyltransferase.

<sup>d</sup>Chr7+/Chr10<sup>-</sup>: chromosome 7 gain and 10 loss.

<sup>e</sup>TERT: telomerase reverse transcriptase.

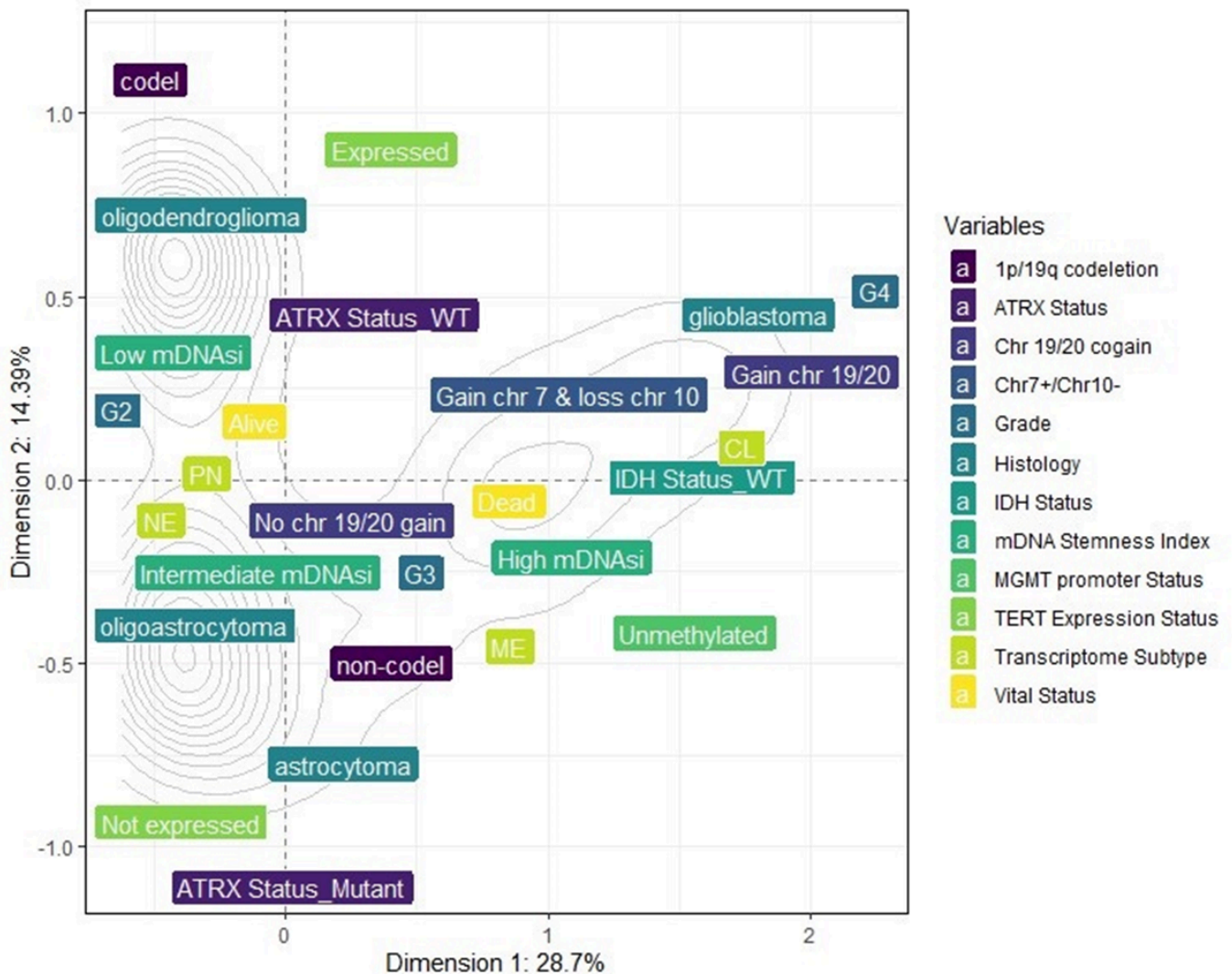
<sup>f</sup>ATRX: Alpha Thalassemia/Mental Retardation Syndrome X-linked.

Using MCA, dimension 1 (x-axis) accounted for 28.7% of the variance, while dimension 2 (y-axis) accounted for 14.39%. The inertia (sum of the variances) for these 2 dimensions was 43.09%. The variance of the overall dimensions (18 dimensions) for the combinations of the variables is illustrated in [Multimedia Appendix 18](#). The total inertia (sum of the variances) was 1.5. The 2D perceptual map exhibited the associations between the categories of each variable ([Figure 3](#)). The perceptual map reveals categories such as GBM, unmethylated MGMT promoter, IDH wild type, Chr7 gain and Chr10 loss, grade 4, GBM ATRX wild

type, TERT expression, non-codel 1p.19q, and CL and ME transcriptome subtypes are closely associated with high mDNAsi, appearing along the positive x-axis (dimension 1). Conversely, categories like oligoastrocytomas and oligodendrogliomas, grade 2, IDH mutant, codel 1p19q, methylated MGMT promoter, no combined copy number alterations, no expression of TERT, ATRX mutant, and PN and NE transcriptome subtypes are closely associated with alive vital status, appearing along the negative x-axis (dimension 1) ([Figure 3](#)).



**Figure 3.** Multiple correspondence analysis (MCA) 2D perceptual map demonstrating the association between the categories of each categorical variable. Categories that are closely clustered are strongly associated with each other. Categories such as glioblastoma, unmethylated MGMT promoter, IDH wild type, chromosome 7 gain and 10 loss (Chr7+/Chr10-), grade 4, glioblastoma ATRX wild type, TERT expression, non-codel 1p.19q, and CL and ME transcriptome subtypes are closely associated with high mDNAsi, appearing along the positive x-axis (dimension 1). ATRX: alpha thalassemia/mental retardation syndrome, X-linked; CL: classical; IDH: isocitrate dehydrogenase; mDNAsi: DNA methylation stemness index; ME: mesenchymal; MGMT: methylguanine methyltransferase; NE: neural; PN: proneural; TERT: telomerase reverse transcriptase.



## Discussion

### Principal Findings

Multiple efforts have been made to explore the diversity of oncologic diseases, with significant contributions from genetics, cell and tissue biology, as well as computational and experimental technologies, providing a wealth of information on cancer manifestations. In the field of glioma research, emerging approaches have sought to clarify tumor pathology and grading through the introduction of novel types and subtypes, as well as by identifying molecular markers and genetic mutations that contribute to predicting diagnosis and prognosis. However, it also results in an accumulation of extensive datasets, presenting challenges in interpretation and visualization regarding the associations between prognostic factors. In this study, we used MCA, an unsupervised data science approach, to establish statistical associations between different qualitative variables of gliomas. This method was

able to reduce data dimensionality and represent it on a 2D perceptual map, revealing associations between various established glioma prognostic factors, including histological classification, IDH status, MGMT promoter methylation, and transcriptome subtypes. Furthermore, we associated these clinical and prognostic variables with an epigenetic-based stemness index (mDNAsi), demonstrating that higher stemness levels were associated with poorer prognostic factors, providing a useful tool to associate prognostic markers in brain tumors.

### Comparison to Prior Studies

Several clinical and molecular factors are considered in predicting the prognosis and survival of brain tumors, more specifically for gliomas. Beyond histological classification and tumor grade, genetic and molecular biomarkers have been incorporated as potential prognostic indicators. Thus, we first evaluated the ability of MCA to associate these consolidated prognostic variables with the patient's vital status. Our

findings demonstrate that MCA effectively clusters poor prognostic factors with dead vital status. All these prognostic factors are well consolidated and associated with malignancy of gliomas. IDH mutation represents one of the main prognostic markers for gliomas [25]. It has been identified that one of the mechanisms given by this favorable outcome is the impaired production of nicotinamide adenine dinucleotide phosphate in Krebs cycle caused by IDH1 enzyme mutation that can sensitize tumor cells to chemotherapy and explain the favorable prognosis of patients with IDH mutation [25]. Likewise, co-deletion of 1p19q chromosome arms, especially when combined with other biomarkers such as IDH mutation and TERT expression, has been used as a predictive biomarker and recent studies investigated biological mechanisms to be significantly linked to genes involved in cell division, angiogenesis, and DNA repair responses [26]. Thus, we demonstrated that MCA was able to capture and associate key glioma hallmarks with patients' vital status, which was applied to different clinical variables.

Subsequently, we applied MCA to explore the association between high stemness levels (mDNAsi) and characteristics related to poor prognosis. Stemness has been considered an important phenotype in glioma malignancy and is potentially associated with CL genetic alterations, such as the gain of chromosome 7. Chromosome 7 harbors some key genes related to stemness, including Epidermal Growth Factor Receptor (EGFR), Mesenchymal-Epithelial Transition Factor (MET), and Homeobox A gene (HOXA). A study of 86 GBMs reported that EGFR amplification occurs with higher probability in samples that have a gain of chromosome 7 (82.1%) compared with those without it (66.7%) [27]. In addition, EGFR amplification is more prevalent in IDH-wild-type diffuse gliomas (66.0%) and GBM (85.5%) [28], which are also associated with poorer prognostic factors, consistent with our findings. High mDNAsi has been previously linked to EGFR mutations [21]. The HOXA and MET loci, also located on chromosome 7, have been implicated in stemness-related pathways. Notably, studies have demonstrated interactions between chromosome 7 gain and the expression of a stem cell-related HOX signature in GBMs [29]. Analysis of the MET gene at 7q31.2 revealed that gain occurs in 47% of primary and 44% of secondary GBMs, suggesting that this genetic alteration contributes to the pathogenesis of both GBM subtypes [30].

Overall, relatively few studies have used MCA to explore associations with cancer phenotypes. Previous studies have applied MCA to different approaches, such as analyzing prognosis low rectal cancer surgery [31], investigating the association between some types of cancer in rural or urban areas [15], examining the association between Traditional Chinese Medicine Syndrome and histopathology of colorectal

cancer [32], assessing clinically relevant demographic variables across multiple gastrointestinal cancers [33], and the relationship between types of diagnostic classification in breast cancer [34]. Our study also highlights the utility of MCA in investigating associations within the context of brain tumors. MCA enables the investigation of the pattern among many categorical factors in gliomas, providing a powerful computational approach to identify and test prognostic variables. It was possible to visually and quantitatively represent the associations, which facilitates the identification of distinct patient clusters based on shared prognostic characteristics. Our findings were consistent with previous literature and emphasized stemness as an important phenotype for gliomas.

### Limitations

Our study has inherent limitations. First, as a retrospective analysis of TCGA data, it is subject to selection bias. Second, we associated all the prognostic variables with patients' vital status, which may not be the most optimal variable for determining prognosis. For the future, we intend to improve our model validating its applicability in other prospective datasets. Third, the absence of therapy data is another limitation of this study. Finally, an intrinsic limitation of MCA is that retaining only 2 or 3 dimensions may not sufficiently capture all the significant features in the data. In our analysis, the percentage of explained inertia was approximately 40%. While there is not an accepted threshold for adequately explained inertia, common guidelines recommend retaining dimensions that represent over 70% of the inertia [35]. However, explained inertia in the range of 40%-60% is often considered informative, and the interpretability and relevance of the patterns revealed by the dimensions are frequently more important than the exact percentage of inertia explained, especially in a complex heterogeneous disease such as brain tumors [36].

### Conclusion and Future Perspectives

In conclusion, our findings suggest that MCA is a valuable tool for understanding the interdependence between prognostic markers in gliomas. MCA facilitates the exploration of a large-scale dataset and enhances the identification of associations. Considering the advances in computational oncology and the emergence of new oncological features, such as stemness phenotype, incorporating MCA into cancer research as an approach to exploring the complex heterogeneity of the oncologic field becomes a powerful tool for simplifying data management. It contributes to researchers statistically identifying associations between variables within extensive databases and improves the visual representation, leading to a deeper understanding of cancer findings.

### Acknowledgments

This study has been supported by grants from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and from the Sao Paulo Research Foundation (FAPESP), Brazil (2018/00583-0, 2022/06305-7, 2022/09378-5, 2023/05099-7, 2023/07358-0).

### Data Availability

The datasets generated or analyzed during this study are available at National Institutes of Health Genomic Data Commons (GDC) [37]. The workflow to generate the DNA methylation stemness index (mDNAsI) can be accessed at GitHub [38].

---

### Authors' Contributions

MEGJ conducted the study, contributing to the acquisition of data, data analysis and interpretation, production of tables and figures, and wrote the first version of the manuscript. HF contributed to the interpretation and discussion of data and corrected the final version of the manuscript. TMM contributed to the acquisition, interpretation, and discussion of data, and corrected the final version of the manuscript. PLPX contributed to the concept and design of the study, data analysis and interpretation, funding, and corrected the final version of the manuscript.

---

### Conflicts of Interest

None declared.

---

### Multimedia Appendix 1

Individual contingency tables for cancer type.

[[XLSX File \(Microsoft Excel File\), 29 KB-Multimedia Appendix 1](#)]

---

### Multimedia Appendix 2

Individual contingency tables for histology.

[[XLSX File \(Microsoft Excel File\), 32 KB-Multimedia Appendix 2](#)]

---

### Multimedia Appendix 3

Individual contingency tables for grade.

[[XLSX File \(Microsoft Excel File\), 28 KB-Multimedia Appendix 3](#)]

---

### Multimedia Appendix 4

Individual contingency tables for gender.

[[XLSX File \(Microsoft Excel File\), 25 KB-Multimedia Appendix 4](#)]

---

### Multimedia Appendix 5

Individual contingency tables for vital status.

[[XLSX File \(Microsoft Excel File\), 23 KB-Multimedia Appendix 5](#)]

---

### Multimedia Appendix 6

Individual contingency tables for IDH (isocitrate dehydrogenase) status.

[[XLSX File \(Microsoft Excel File\), 22 KB-Multimedia Appendix 6](#)]

---

### Multimedia Appendix 7

Individual contingency tables for X1p.19q.codeletion.

[[XLSX File \(Microsoft Excel File\), 20 KB-Multimedia Appendix 7](#)]

---

### Multimedia Appendix 8

Individual contingency tables for MGMT (methylguanine methyltransferase) promoter.

[[XLSX File \(Microsoft Excel File\), 19 KB-Multimedia Appendix 8](#)]

---

### Multimedia Appendix 9

Individual contingency tables for Chr 7 gain and Chr 10 loss.

[[XLSX File \(Microsoft Excel File\), 17 KB-Multimedia Appendix 9](#)]

---

### Multimedia Appendix 10

Individual contingency tables for Chr 19/20 co-gain.

[[XLSX File \(Microsoft Excel File\), 16 KB-Multimedia Appendix 10](#)]

---

### Multimedia Appendix 11

Individual contingency tables for TERT (telomerase reverse transcriptase) expression status.

[[XLSX File \(Microsoft Excel File\), 14 KB-Multimedia Appendix 11](#)]

---

### Multimedia Appendix 12

Individual contingency tables for ATRX (Alpha Thalassemia/Mental Retardation Syndrome X-linked alpha thalassemia/mental retardation syndrome, X-linked) status.

[\[XLSX File \(Microsoft Excel File\), 12 KB-Multimedia Appendix 12\]](#)

### Multimedia Appendix 13

Individual contingency tables for DAXX status.

[\[XLSX File \(Microsoft Excel File\), 10 KB-Multimedia Appendix 13\]](#)

### Multimedia Appendix 14

Fisher exact test and  $\chi^2$  test for vital status  $\times$  glioma prognostic factors.

[\[XLSX File \(Microsoft Excel File\), 25 KB-Multimedia Appendix 14\]](#)

### Multimedia Appendix 15

Percentage of explained variances of the overall (17) dimensions.

[\[PNG File \(Portable Network Graphics File\), 162 KB-Multimedia Appendix 15\]](#)

### Multimedia Appendix 16

Individual contingency table for mDNAsi.

[\[XLSX File \(Microsoft Excel File\), 31 KB-Multimedia Appendix 16\]](#)

### Multimedia Appendix 17

Fisher exact test and  $\chi^2$  test for mDNAsi (DNA methylation stemness index)  $\times$  glioma prognostic factors.

[\[XLSX File \(Microsoft Excel File\), 23 KB-Multimedia Appendix 17\]](#)

### Multimedia Appendix 18

Percentage of explained variances of the overall (18) dimensions.

[\[PNG File \(Portable Network Graphics File\), 10 KB-Multimedia Appendix 18\]](#)

## References

1. Hanahan D. Hallmarks of cancer: new dimensions. *Cancer Discov.* Jan 2022;12(1):31-46. [doi: [10.1158/2159-8290.CD-21-1059](https://doi.org/10.1158/2159-8290.CD-21-1059)] [Medline: [35022204](https://pubmed.ncbi.nlm.nih.gov/35022204/)]
2. Dagogo-Jack I, Shaw AT. Tumour heterogeneity and resistance to cancer therapies. *Nat Rev Clin Oncol.* Feb 2018;15(2):81-94. [doi: [10.1038/nrclinonc.2017.166](https://doi.org/10.1038/nrclinonc.2017.166)] [Medline: [29115304](https://pubmed.ncbi.nlm.nih.gov/29115304/)]
3. Brierley J, O'Sullivan B, Asamura H, et al. Global consultation on cancer staging: promoting consistent understanding and use. *Nat Rev Clin Oncol.* Dec 2019;16(12):763-771. [doi: [10.1038/s41571-019-0253-x](https://doi.org/10.1038/s41571-019-0253-x)] [Medline: [31388125](https://pubmed.ncbi.nlm.nih.gov/31388125/)]
4. Weller M, Wick W, Aldape K, et al. Glioma. *Nat Rev Dis Primers.* Jul 16, 2015;1:15017. [doi: [10.1038/nrdp.2015.17](https://doi.org/10.1038/nrdp.2015.17)] [Medline: [27188790](https://pubmed.ncbi.nlm.nih.gov/27188790/)]
5. Louis DN, Ohgaki H, Wiestler OD, et al. The 2007 WHO classification of tumours of the central nervous system. *Acta Neuropathol.* Aug 2007;114(2):97-109. [doi: [10.1007/s00401-007-0243-4](https://doi.org/10.1007/s00401-007-0243-4)] [Medline: [17618441](https://pubmed.ncbi.nlm.nih.gov/17618441/)]
6. Louis DN, Perry A, Wesseling P, et al. The 2021 WHO classification of tumors of the central nervous system: a summary. *Neuro Oncol.* Aug 2, 2021;23(8):1231-1251. [doi: [10.1093/neuonc/noab106](https://doi.org/10.1093/neuonc/noab106)] [Medline: [34185076](https://pubmed.ncbi.nlm.nih.gov/34185076/)]
7. Ayob AZ, Ramasamy TS. Cancer stem cells as key drivers of tumour progression. *J Biomed Sci.* Mar 6, 2018;25(1):20. [doi: [10.1186/s12929-018-0426-4](https://doi.org/10.1186/s12929-018-0426-4)] [Medline: [29506506](https://pubmed.ncbi.nlm.nih.gov/29506506/)]
8. Battle E, Clevers H. Cancer stem cells revisited. *Nat Med.* Oct 6, 2017;23(10):1124-1134. [doi: [10.1038/nm.4409](https://doi.org/10.1038/nm.4409)] [Medline: [28985214](https://pubmed.ncbi.nlm.nih.gov/28985214/)]
9. Wang Q, Hu B, Hu X, et al. Tumor evolution of glioma-intrinsic gene expression subtypes associates with immunological changes in the microenvironment. *Cancer Cell.* Jul 10, 2017;32(1):42-56. [doi: [10.1016/j.ccell.2017.06.003](https://doi.org/10.1016/j.ccell.2017.06.003)] [Medline: [28697342](https://pubmed.ncbi.nlm.nih.gov/28697342/)]
10. Ortensi B, Setti M, Osti D, Pelicci G. Cancer stem cell contribution to glioblastoma invasiveness. *Stem Cell Res Ther.* Feb 28, 2013;4(1):18. [doi: [10.1186/scrt166](https://doi.org/10.1186/scrt166)] [Medline: [23510696](https://pubmed.ncbi.nlm.nih.gov/23510696/)]
11. Tan J, Zhu H, Tang G, et al. Molecular subtypes based on the stemness index predict prognosis in glioma patients. *Front Genet.* 2021;12:616507. [doi: [10.3389/fgene.2021.616507](https://doi.org/10.3389/fgene.2021.616507)] [Medline: [33732284](https://pubmed.ncbi.nlm.nih.gov/33732284/)]
12. Sourial N, Wolfson C, Zhu B, et al. Correspondence analysis is a useful tool to uncover the relationships among categorical variables. *J Clin Epidemiol.* Jun 2010;63(6):638-646. [doi: [10.1016/j.jclinepi.2009.08.008](https://doi.org/10.1016/j.jclinepi.2009.08.008)] [Medline: [19896800](https://pubmed.ncbi.nlm.nih.gov/19896800/)]
13. Li BH, Sun ZQ, Dong SF. Correspondence analysis and its application in oncology. *Commun Stat Theory Methods.* Mar 19, 2010;39(7):1229-1236. [doi: [10.1080/03610920902871446](https://doi.org/10.1080/03610920902871446)]

14. Costa PS, Santos NC, Cunha P, Cotter J, Sousa N. The use of multiple correspondence analysis to explore associations between categories of qualitative variables in healthy ageing. *J Aging Res.* 2013;2013:302163. [doi: [10.1155/2013/302163](https://doi.org/10.1155/2013/302163)] [Medline: [24222852](https://pubmed.ncbi.nlm.nih.gov/24222852/)]
15. Florensa D, Godoy P, Mateo J, et al. The use of multiple correspondence analysis to explore associations between categories of qualitative variables and cancer incidence. *IEEE J Biomed Health Inform.* Sep 2021;25(9):3659-3667. [doi: [10.1109/JBHI.2021.3073605](https://doi.org/10.1109/JBHI.2021.3073605)] [Medline: [33857006](https://pubmed.ncbi.nlm.nih.gov/33857006/)]
16. van Horn A, Weitz CA, Olszowy KM, et al. Using multiple correspondence analysis to identify behaviour patterns associated with overweight and obesity in Vanuatu adults. *Public Health Nutr.* Jun 2019;22(9):1533-1544. [doi: [10.1017/S1368980019000302](https://doi.org/10.1017/S1368980019000302)] [Medline: [30846019](https://pubmed.ncbi.nlm.nih.gov/30846019/)]
17. Śledzińska P, Bebyn MG, Furtak J, Kowalewski J, Lewandowska MA. Prognostic and predictive biomarkers in gliomas. *Int J Mol Sci.* Sep 26, 2021;22(19):10373. [doi: [10.3390/ijms221910373](https://doi.org/10.3390/ijms221910373)] [Medline: [34638714](https://pubmed.ncbi.nlm.nih.gov/34638714/)]
18. Sokolov A, Paull EO, Stuart JM. ONE-class detection of cell states in tumor subtypes. Presented at: Proceedings of the Pacific Symposium; Jan 4-8, 2016; Kohala Coast, Hawaii, USA. [doi: [10.1142/9789814749411\\_0037](https://doi.org/10.1142/9789814749411_0037)]
19. Salomonis N, Dexheimer PJ, Omberg L, et al. Integrated genomic analysis of diverse induced pluripotent stem cells from the progenitor cell biology consortium. *Stem Cell Rep.* Jul 12, 2016;7(1):110-125. [doi: [10.1016/j.stemcr.2016.05.006](https://doi.org/10.1016/j.stemcr.2016.05.006)] [Medline: [27293150](https://pubmed.ncbi.nlm.nih.gov/27293150/)]
20. Daily K, Ho Sui SJ, Schriml LM, et al. Molecular, phenotypic, and sample-associated data to describe pluripotent stem cell lines and derivatives. *Sci Data.* Mar 28, 2017;4:170030. [doi: [10.1038/sdata.2017.30](https://doi.org/10.1038/sdata.2017.30)] [Medline: [28350385](https://pubmed.ncbi.nlm.nih.gov/28350385/)]
21. Malta TM, Sokolov A, Gentles AJ, et al. Machine learning identifies stemness features associated with oncogenic dedifferentiation. *Cell.* Apr 5, 2018;173(2):338-354. [doi: [10.1016/j.cell.2018.03.034](https://doi.org/10.1016/j.cell.2018.03.034)] [Medline: [29625051](https://pubmed.ncbi.nlm.nih.gov/29625051/)]
22. Lê S, Josse J, Husson F. FactoMineR: an R package for multivariate analysis. *J Stat Softw.* Mar 2008;25(1):1-18. [doi: [10.18637/jss.v025.i01](https://doi.org/10.18637/jss.v025.i01)] [Medline: [27348562](https://pubmed.ncbi.nlm.nih.gov/27348562/)]
23. The Cancer Genome Atlas program (TCGA). Center for Cancer Genomics. URL: <https://www.cancer.gov/tcga> [Accessed 2025-03-06]
24. The Cancer Genome Atlas program. National Cancer Institute. URL: <https://www.cancer.gov/ccg/research/structural-genomics/tcga/history/policies/tcga-human-subjects-data-policies.pdf> [Accessed 2025-03-06]
25. Bleeker FE, Atai NA, Lamba S, et al. The prognostic IDH1( R132 ) mutation is associated with reduced NADP+-dependent IDH activity in glioblastoma. *Acta Neuropathol.* Apr 2010;119(4):487-494. [doi: [10.1007/s00401-010-0645-6](https://doi.org/10.1007/s00401-010-0645-6)] [Medline: [20127344](https://pubmed.ncbi.nlm.nih.gov/20127344/)]
26. Chai RC, Zhang KN, Chang YZ, et al. Systematically characterize the clinical and biological significances of 1p19q genes in 1p/19q non-codeletion glioma. *Carcinogenesis.* Oct 16, 2019;40(10):1229-1239. [doi: [10.1093/carcin/bgz102](https://doi.org/10.1093/carcin/bgz102)] [Medline: [31157866](https://pubmed.ncbi.nlm.nih.gov/31157866/)]
27. McNulty SN, Cottrell CE, Vigh-Conrad KA, et al. Beyond sequence variation: assessment of copy number variation in adult glioblastoma through targeted tumor somatic profiling. *Hum Pathol.* Apr 2019;86:170-181. [doi: [10.1016/j.humpath.2018.12.004](https://doi.org/10.1016/j.humpath.2018.12.004)] [Medline: [30594748](https://pubmed.ncbi.nlm.nih.gov/30594748/)]
28. Wang H, Zhang X, Liu J, et al. Clinical roles of EGFR amplification in diffuse gliomas: a real-world study using the 2021 WHO classification of CNS tumors. *Front Neurosci.* 2024;18:1308627. [doi: [10.3389/fnins.2024.1308627](https://doi.org/10.3389/fnins.2024.1308627)] [Medline: [38595969](https://pubmed.ncbi.nlm.nih.gov/38595969/)]
29. Kurscheid S, Bady P, Sciuscio D, et al. Chromosome 7 gain and DNA hypermethylation at the HOXA10 locus are associated with expression of a stem cell related HOX-signature in glioblastoma. *Genome Biol.* Jan 27, 2015;16(1):16. [doi: [10.1186/s13059-015-0583-7](https://doi.org/10.1186/s13059-015-0583-7)] [Medline: [25622821](https://pubmed.ncbi.nlm.nih.gov/25622821/)]
30. Pierscianek D, Kim YH, Motomura K, et al. MET gain in diffuse astrocytomas is associated with poorer outcome. *Brain Pathol.* Jan 2013;23(1):13-18. [doi: [10.1111/j.1750-3639.2012.00609.x](https://doi.org/10.1111/j.1750-3639.2012.00609.x)] [Medline: [22672415](https://pubmed.ncbi.nlm.nih.gov/22672415/)]
31. Mancini R, Pattaro G, Diodoro MG, et al. Tumor regression grade after neoadjuvant chemoradiation and surgery for low rectal cancer evaluated by multiple correspondence analysis: ten years as minimum follow-up. *Clin Colorectal Cancer.* Mar 2018;17(1):e13-e19. [doi: [10.1016/j.clcc.2017.06.004](https://doi.org/10.1016/j.clcc.2017.06.004)] [Medline: [28865674](https://pubmed.ncbi.nlm.nih.gov/28865674/)]
32. Wu T, Zhang S, Guo S, et al. Correspondence analysis between traditional Chinese medicine (TCM) syndrome differentiation and histopathology in colorectal cancer. *Eur J Integr Med.* Aug 2015;7(4):342-347. [doi: [10.1016/j.eujim.2015.07.003](https://doi.org/10.1016/j.eujim.2015.07.003)]
33. Kramer RJ, Rhodin KE, Therien A, et al. Unsupervised clustering using multiple correspondence analysis reveals clinically-relevant demographic variables across multiple gastrointestinal cancers. *Surgical Oncology Insight.* Mar 2024;1(1):100009. [doi: [10.1016/j.soi.2024.100009](https://doi.org/10.1016/j.soi.2024.100009)]
34. Nadjib Bustan M, Arif Tiro M, Annas S. Correspondence analysis of breast cancer diagnosis classification. *J Phys Conf Ser.* Jun 1, 2019;1244(1):012030. [doi: [10.1088/1742-6596/1244/1/012030](https://doi.org/10.1088/1742-6596/1244/1/012030)]

35. Higgs NT. Practical and innovative uses of correspondence analysis. *R Stat Soc Ser D (The Statistician)*. 1991;40(2):183. [doi: [10.2307/2348490](https://doi.org/10.2307/2348490)]
36. Husson F, Lê S, Pagès J. *Exploratory Multivariate Analysis by Example Using R*. CRC Press; 2011. ISBN: 978-1439826454
37. Machine learning identifies stemness features associated with oncogenic dedifferentiation. National Cancer Institute. URL: <https://gdc.cancer.gov/about-data/publications/PanCanStemness-2018> [Accessed 2025-03-06]
38. PanCanStem: reproducing mrnasi from PMID: 29625051. GitHub. URL: <https://github.com/ArtemSokolov/PanCanStem> [Accessed 2025-03-06]

## Abbreviations

**ASR:** adjusted standardized residual  
**ATRX:** alpha thalassemia/mental retardation syndrome, X-linked  
**Chr:** chromosome  
**Chr7+/Chr10-:** chromosome 7 gain and 10 loss  
**CL:** classical  
**CSC:** cancer stem cell  
**GMB:** glioblastoma  
**IDH:** isocitrate dehydrogenase  
**MCA:** multiple correspondence analysis  
**mDNasi:** DNA methylation stemness index  
**ME:** mesenchymal  
**MGMT:** methylguanine methyltransferase  
**NE:** neural  
**PR:** proneural  
**TCGA:** the Cancer Genome Atlas  
**TERT:** telomerase reverse transcriptase

*Edited by Ece Uzun; peer-reviewed by Chao Tang, Jiaying Lai; submitted 21.08.2024; final revised version received 22.11.2024; accepted 04.02.2025; published 12.03.2025*

*Please cite as:*

*Goes Job ME, Fukumasu H, Malta TM, Porfirio Xavier PL*

*Investigating Associations Between Prognostic Factors in Gliomas: Unsupervised Multiple Correspondence Analysis*

*JMIR Bioinform Biotech 2025;6:e65645*

URL: <https://bioinform.jmir.org/2025/1/e65645>

doi: [10.2196/65645](https://doi.org/10.2196/65645)

© Maria Eduarda Goes Job, Heidge Fukumasu, Tathiane Maistro Malta, Pedro Luiz Porfirio Xavier. Originally published in *JMIR Bioinformatics and Biotechnology* (<https://bioinform.jmir.org>), 12.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Bioinformatics and Biotechnology*, is properly cited. The complete bibliographic information, a link to the original publication on <https://bioinform.jmir.org/>, as well as this copyright and license information must be included.