Radiomic and Genomic Approaches to Survival Stratification in Adult-type Diffuse Glioma

Nicholas Nuechterlein

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2023

Reading Committee:

Linda Shapiro, Chair

Eric C. Holland

Patrick J. Cimino

Program Authorized to Offer Degree:

Paul G. Allen School of Computer Science & Engineering

University of Washington

Abstract

Radiomic and Genomic Approaches to Survival Stratification in Adult-type Diffuse Glioma

Nicholas Nuechterlein

Chair of the Supervisory Committee:
Linda Shapiro
Paul G. Allen School of Computer Science & Engineering

Adult-type diffuse gliomas are the most common primary adult malignant brain tumors and are generally considered incurable. Alongside tumor molecular heterogeneity and barriers to drug delivery, inadequate survival stratification is a critical challenge for patient clinical management, drug development, and clinical trial enrollment. In this work, we will discuss several radiomic and genomic approaches to risk stratification of types of adult-type diffuse glioma using magnetic resonance imaging (MRI), somatic copy number alteration (SCNA), gene expression, DNA methylation, and clinical data. From MRI data, we train an efficient adult-type diffuse glioma tumor segmentation model that automates a critical piece of standard radiomic pipelines, present a novel radiomic feature selection method that allows for accurate prediction of survival-associated subgroups of IDH-wildtype glioblastoma, and showcase the potential utility of multitask learning and unlabeled MRI data on adult-type diffuse glioma classification and survival prediction tasks. From genomic data, we assemble a large SCNA dataset and train models to predict the mutational status of genes *IDH1* and *IDH2* mutation and the codeletion status of chromosome arms 1p and 19q in order to assign molecular diagnoses to patients with antiquated SCNA data and unknown diagnoses. We also nominate an oncogene that may characterize an aggressive subgroup of IDH-mutant and 1p/19q-codeleted oligodendroglioma associated with, but independently prognostic of, patient age. Altogether, our work in risk-stratifying adult-type diffuse glioma has implica-

tions for upfront clinical trial recommendations for patients diagnosed with IDH-wildtype glioblastoma, treatment toxicity calibration for risk-assessed oligodendroglioma, and clinical trial arm sorting, which may accelerate the approval of new effective therapies.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

ADULT-TYPE DIFFUSE GLIOMA: Adult-type diffuse glioma make up 80% of all primary adult malignant brain tumors and are generally considered incurable.

BIOMARKER: Biomarkers are any measurable characteristic of a disease, including clinical features such as age, radiographic features such as tumor volume, histopathologic features such as necrosis, and genomic alterations such as IDH mutations and 1p/19q-codeletions.

IDH MUTATION: IDH mutations are specific point mutations in the genes *IDH1* and *IDH2* that lead IDH-mutant tumors to have significantly different chromatin structures compared to IDH-wildtype glioblastoma. IDH enzymes catalyze the conversion of isocitrate to alpha-ketoglutarate (α-KG), which induces the production of D-2-hydroxyglutarate (2-HG), which in turn inhibits histone and DNA demethylases.

1P/19Q CO-DELETION: In adult-diffuse glioma, the 1p/19q-codeletion is a translocation-mediated event that results in single-copy loss of chromosome arms 1p and 19q and intact status of chromosome arms 1q and 19p. The presence or absence of a 1p/19q-codeletion divides IDH-mutant adult-type diffuse gliomas into oligodendroglioma, which are 1p/19q-codeleted, and IDH-mutant astrocytoma, which are not.

MRI: Magnetic resonance imaging (MRI) is a non-invasive medical imaging technique that uses magnetic fields to produce detailed images of the body's interior.

PERITUMORAL EDEMA: Peritumoral edema refers to swelling that occurs in the tissues surrounding a tumor and is clearly visible on T2-weighted MRI sequences, which are sensitive to increased water content. Peritumoral edema is a common feature in brain tumors, particularly in malignant gliomas.

SCNA: Somatic copy number alterations (SCNAs) are losses or gains of particular genes or groups of genes in the genome. The loss of one or both copies of a gene that functions as a tumor suppressor, or the duplication or amplification of an oncogene, can cause oncogenesis or increase tumor aggressiveness.

GENE EXPRESSION: Gene expression data quantifies the level to which a gene's DNA sequence is translated into RNA. Gene expression data is commonly acquired via bulk RNAseq or single-cell RNAseq.

DNA METHYLATION: DNA methylation is an epigenetic process that can change the activity of a gene without changing the gene's DNA sequence. DNA methylation typically occurs in the promoter region of genes, where it can play a role in regulating gene expression, but can also occur in the gene body.

RADIOMICS: Radiomics is a field in medical imaging that employs machine learning to associate quantitative imaging with underlying characteristics of the imaged tissue.

BIOINFORMATICS: Bioinformatics is a field that applies principles of biology, data science, statistics, and machine learning to understand the complex nature of biological entities.

TCGA: The Cancer Genome Atlas (TCGA) is a publicly available cancer genomics dataset that has characterized over 20,000 patients across 33 cancer types. The TCGA generated genomic data for overy 1,000 adult-type diffuse glioma.

BRATS: The Multimodal Brain Tumor Segmentation Challenge (BraTS) provides preoperative multimodal MRI scans of adult-type diffuse gliomas and corresponding segmentation masks delineating multiple tumor compartments.

ACKNOWLEDGMENTS

I am deeply grateful for those who have contributed to my work and well-being over the past five years. Goodness knows, I needed all of you. I have had the rare opportunity to work with a trio of generous mentors who have guided me with grace. To Professor Linda Shapiro, thank you for accepting me into your lab, and for supporting and advocating for me and all of your students whose interests you often put before your own. You have always been available, reliable, encouraging, and caring. To Dr. Eric Holland, thank you for taking a chance on a stranger you met on an airplane. Your vision, kindness, and willingness to collaborate were pivotal to the second half of my graduate career and essential for my graduation. To Dr. PJ Cimino, thank you for your patience, prudence, and push for productivity. You have made research not only interesting but fun. You taught me what it means to ask a question, how to write for reviewers, and how to keep a level head. I also must thank Dr. James Fink for his help in introducing me to neuroradiology, and Dr. Lynne Taylor for being a symbol of neuro-oncology clinical practice and patient advocacy at its best.

I am also indebted to close collaborators and friends Dr. Beibin Li, Sonali Arora, Dr. Siohban Pattwell, Saygin Seyfioglu, and Dr. Sachin Mehta. Beibin, thank you for the long conversations over coffee and in the backcountry, and Sonali for your bioinformatic advice. To my labmates Dr. Shima Noffalah, Rashmi Muddurulu, Dr. Deepali Aneja, Fatemeh Ghezloo, Mahtab Bigverdi, Wenjun Wu, and Wisdom Ikezogwo: thank you for the campus walks, afternoon conversations, and your rich perspectives. Life is all that much lighter when you have kind people around you. To Josh Gardner, Hannah Lee, and Professor Jamie Morgenstern – thank you for making teaching my last, best experience in graduate school. And to my dear friends that I have not already mentioned–Philip, Taylor, Colin, Sam, Grace,

Allan, Cale, Henry, Michelle, Hillary, Frank, Peter, Gus, and Erin–thank you for the joy and meaning you have brought to my life.

To my family—Mom, Dad, Noah, Kai—I could not ask for better. Those weeks we have spent together are some of the best I can remember. May my life take me closer to you.

And to Danielle—my partner, my wife, and my best friend—thank you for making my life such a joy and for making these five years, which should have been stress-ridden, feel easy. You have made my everyday new and interesting. You are intertwined in anything I accomplish. It is not me, it is us.

And, last, to the patients who have been taken by these diseases, and to those who are living with these diseases, and to their caretakers and their families, to those I have known and those I have not: you are the reason for this research. Your courage, your strength, your fears, your every bad day, and your hope—all that makes you—are an absolute inspiration and why every paper counts. May we all join together and find a cure.

DEDICATION

to my dear wife, Danielle

Chapter 1

INTRODUCTION

Adult-type diffuse gliomas make up 80% of all primary adult malignant brain tumors and are generally considered incurable [1, 2]. While patient outcomes have improved dramatically for many other cancer types over the last decade, adult-type diffuse glioma outcomes have remained static primarily due to unique biological barriers in the brain [3, 4]. However, research on biomarkers, any measurable characteristics of a tumor, has led to significant progress in adult-type diffuse glioma sub-classification. Molecular biomarkers have established three rigorously defined types of adult-type diffuse glioma, each with a distinct prognosis, demographic, and biological profile [5, 6]. Still, patient outcomes within tumor type remain poorly stratified, which hampers clinical management and drug development. Prognostic biomarkers capable of accurately risk-stratifying patients within each tumor type have the potential to improve clinical trials, expand treatment options for patients with the grimmest prognoses, and allow favorable risk patients to safely pursue treatment regimens with less toxicity. Thus, survival-associated biomarker discovery is of utmost importance in adult-type diffuse glioma.

The 2016 World Health Organization (WHO) classification of central nervous system (CNS) tumors, and its refined 2021 update, integrated molecular biomarkers into diagnostic criteria, which improved diagnostic and prognostic accuracy [5, 6]. These classifications consolidated adult-type diffuse glioma into three survival stratifying categories based on the mutational status of genes *IDH1* and *IDH2* and the co-deletion status of whole chromosome arms 1p and 19q (Figure 1.1A,B). The median survival is 15 months for patients diagnosed with IDH-wildtype glioblastoma, seven years for those diagnosed with IDH-mutant astrocytoma, and ten years for those diagnosed with IDH-mutant and 1p/19q co-deleted

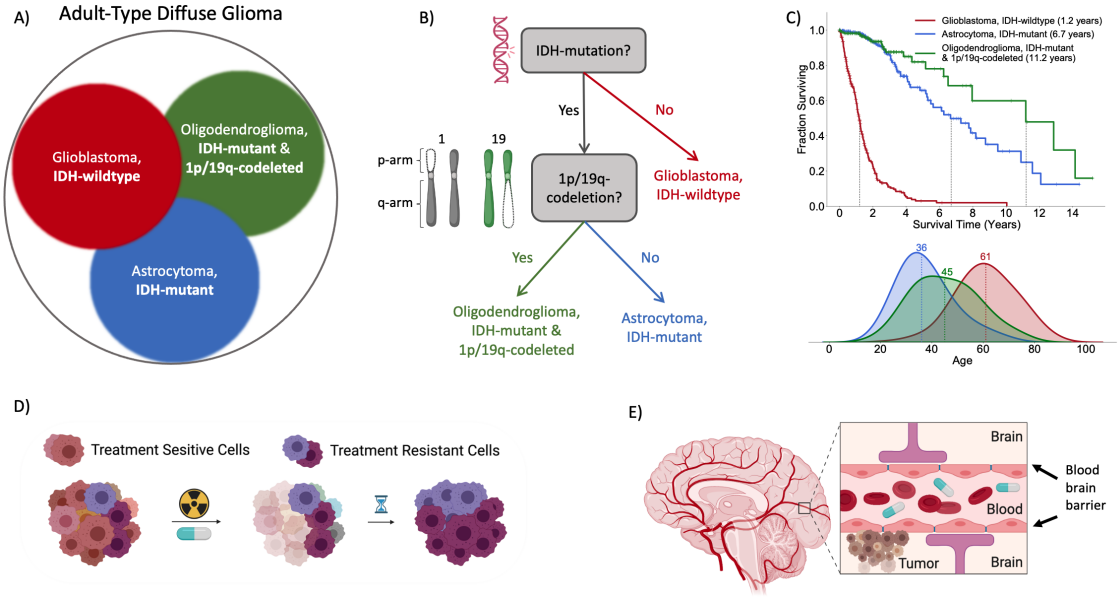


Figure 1.1: **A)** Adult-type diffuse glioma consists of IDH-wildtype glioblastoma, IDH-mutant astrocytoma, and 1p/19q-codeleted oligodendroglioma. **B)** The WHO 2021 molecular classification of adult-type diffuse glioma based on IDH mutation and 1p/19q-codeletion status. A 1p/19q-codeletion is illustrated. **C)** Kaplan-Meier survival curves and patient age density plots for adult-type diffuse glioma patients in The Cancer Genome Atlas (TCGA). **D)** Treatment-resistant cellular sub-populations of heterogeneous adult-type diffuse glioma tumors persist through radiation and chemotherapy and proceed to replicate and drive tumor recurrence. **E)** The blood-brain barrier complicates glioma treatment because it can prevent therapies from reaching tumor cells, masking their potential efficacy.

oligodendroglioma [7] (Figure 1.1C). The average age at diagnosis is 61 for IDH-wildtype glioblastoma, 36 for IDH-mutant astrocytoma, and 45 for oligodendroglioma. However, despite differences in molecular profile, prognosis, and age distribution, all types of adult-type diffuse glioma receive the same standard-of-care treatment which nearly uniformly fails, albeit at different paces [8, 9].

The primary clinical challenges for treating adult-type diffuse glioma are developing ther-

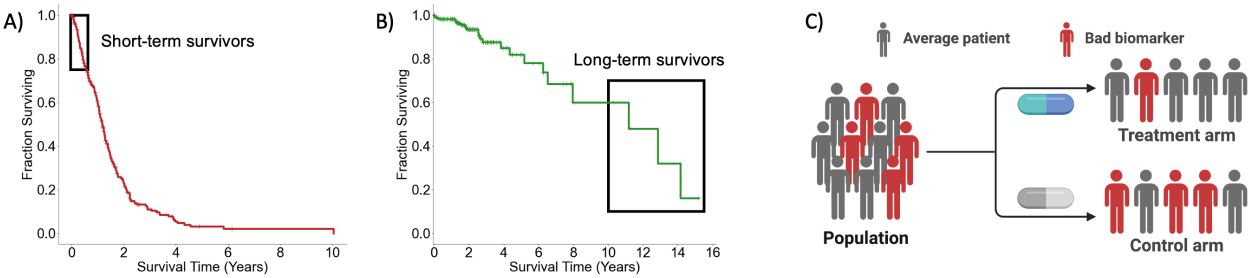


Figure 1.2: **A)** IDH-wildtype glioblastoma short-term survivors who would likely benefit from upfront clinical trials. **B)** Oligodendroglioma long-term survivors who may be candidates for mild upfront therapy. **C)** Illustration of an unbalanced distribution of negative biomarkers across Phase 3 clinical trial control and treatment arms.

apies that are robust to tumor molecular heterogeneity and designing drug delivery methods that can cross the blood-brain barrier and reach all tumor cells [10, 11]. Even in patients who seem to respond to therapy, small treatment-resistant cell subpopulations survive and eventually drive tumor recurrence (Figure 1.1D). Moreover, independent of treatment-resistant cell subpopulations, the blood-brain barrier, which restricts the passage of large molecules into brain tissue, makes it difficult for pharmacological agents, such as chemotherapy, to reach tumor cells (Figure 1.1E). These obstacles both demand the attention of more clinical trials and explain why so many have failed.

Clinical trial design and patients seeking to enroll in clinical trials stand to benefit from the discovery of biomarkers capable of stratifying survival [12, 6, 13, 14]. Identification of extremely short-term survivors at the time of diagnosis could prompt upfront clinical trial enrollment, which would provide novel treatments to the patients most in need and accelerate drug development [15, 16]. Currently, most clinical trials are offered to patients once standard-of-care radiation and chemotherapy have failed, which excludes patients burdened with extremely aggressive IDH-wildtype glioblastoma tumors from enrolling in clinical trials because they are often too ill to qualify by the time standard-of-care treatment is abandoned (Figure 1.2A). On the opposite end of the survival spectrum, ostensibly long-term survivors,

such as patients diagnosed with oligodendroglioma, may elect to temper or delay treatment if they have favorable prognostic markers, thereby avoiding unnecessary treatment-related toxicity (Figure 1.2B).

More generally, knowledge of prognostic biomarkers can improve the assessment of survival benefit in clinical trials, better balance patient prognoses across treatment and control arms, and aid the interpretation of past clinical trial results. Perceived therapy-induced survival benefit in Phase 2 trials, which test treatment efficacy, may be due to fewer poor prognostic biomarkers in enrolled patients, especially if these biomarkers are underrepresented in the treatment arm (Figure 1.2C). These misleading results can cause ineffective therapies to pass Phase 2 trials and fail in expensive, time-consuming, multi-institutional Phase 3 trials where randomization and larger sample sizes better balance biomarker distributions. Knowledge of prognostic biomarkers could calibrate Phase 2 survival expectations and balance trial arms. Results of previous clinical trials can also be reinspected with knowledge of prognostic biomarkers to reassess the difference between the observed outcomes and expected outcomes, and to determine whether a therapy was beneficial for patients with certain biomarkers and not for others [17].

Biomarkers in adult-type diffuse glioma can be derived from magnetic resonance imaging (MRI) data, histopathologic data, and genomic data. MRI is non-invasive and captures rich, volumetric images of the entire tumor, histopathologic images show microscopic cellular features, and genomic data can determine a host of molecular characteristics—including DNA alterations, downstream gene expression, and epigenetic changes—at various levels of resolution. MRI-based biomarkers are attractive because MRI is safe, repeatable, and relatively affordable, however, these biomarkers are difficult to reliably reproduce and validate biologically. Molecular biomarkers derived from genomic data, although dependent on invasive surgery, are routinely collected, because surgical resection is standard-of-care. Molecular biomarkers are usually unambiguous, robustly validated in animal models, and reflective of a tumor's biology. Histological biomarkers remain relevant but have seeded much of their utility to molecular markers [5, 6, 12, 18, 19]. Digital pathology analyses in the interopera-

tive setting have demonstrated significant clinical utility but are generally superseded once the status of established genomic biomarkers is ascertained [20, 21, 22]. Thus, we focused our work on analyses of MRI data and genomic data, building on the fields of radiomics and applied bioinformatics.

Radiomics is an evolving field in medical imaging that employs machine learning to associate quantitative imaging with underlying characteristics of the imaged tissue [23, 24]. The recent neural network-driven revolution in computer vision has transformed radiomics, replacing handcrafted imaging features with complex, high-dimensional representations of MRI-pictured tumors generated by state-of-the-art computer vision models [25]. These models have been used to predict tumor diagnoses, likely areas of recurrence, treatment response, genomic characteristics, and survival [26, 27, 28, 29, 30, 31, 32, 33]. However, despite excelling in certain adult-type diffuse glioma imaging tasks, such as tumor segmentation, radiomic results are rarely clinically adopted, because they are sensitive to overfitting and thus difficult to validate [34, 35]. Fundamental sources of error in radiomic models, including noise induced by patient movement and differences between scanner manufacturers, are complicated by small, high-dimensional datasets [36]. Work in federated learning, a learning method that allows models to be trained on data from multiple institutions without data exchange, has effectively increased the size of training sets and offers a promising way forward for more robust models [37]. Still, much work remains to standardize data collection protocols, better interpret model results, and create robust lower-dimensional representations of MRI data.

Bioinformatics applies principles of biology, data science, statistics, and machine learning to understand the complex nature of biological entities. Bioinformatic methods have radically changed the landscape of tumor classification by integrating the analyses of multiple genome-wide data modalities. The utilization of unsupervised machine learning methods has uncovered latent genomic biomarkers, usually defined as clusters or genomic signatures, that have led to the discovery, refinement, validation, and prediction of tumor types including adult-type diffuse glioma [38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48]. Broad unsupervised analyses of CNS tumor DNA methylation data have identified previously unknown tumor types

and revised existing diagnoses [47, 49]. Multiple adult-type diffuse glioma unsupervised analyses of genome-wide gene expression and DNA methylation have resulted in distinct clusters that correspond to the IDH-mutation and 1p/19q-codeletion defined types of adult-type diffuse glioma, confirming their status as biologically separate entities [42, 43, 44, 45, 46, 47]. Similar studies also identified a subset of IDH-wildtype tumors no longer considered adult-type [50, 51, 52, 12]. Within molecularly-defined types of adult-type diffuse glioma, studies have purposed further subtyping based on DNA methylation and gene expression signatures, although few have shown prognostic utility [41, 53, 54, 45, 48]. In fact, other than the IDH-mutation and 1p/19q-codeletion, few prognostic biomarkers have been adopted in adult-type diffuse glioma outside of promoter methylation of DNA repair gene MGMT and homozygous deletion of tumor suppressor in *CDKN2A* [55, 46, 56, 18]. Prognostic biomarker discovery in adult-type diffuse glioma requires more research.

We begin this dissertation by detailing our work in radiomics and end with our work in bioinformatics. The outline of this document is as follows. We describe in detail all data used in our work in chapter 2. In chapter 3, we describe an efficient adult-type diffuse glioma tumor segmentation model that automates a critical piece of standard radiomic pipelines [57]. In chapter 4, we present a novel radiomic feature selection method that allows for accurate prediction of survival-associated subgroups of IDH-wildtype glioblastoma from MRI data [16]. In chapter 5, we combine tumor-type prediction with tumor segmentation to showcase the potential utility of multitask learning and unlabeled MRI data on adult-type diffuse glioma classification and survival prediction tasks [58]. In chapter 6, we assemble a large somatic copy number alteration (SCNA) dataset and train models to predict adult-type diffuse glioma molecular diagnoses from antiquated somatic copy number alteration data that was collected before IDH mutation and 1p/19q codeletion statuses were routinely collected [7]. Chapter 7 provides evidence that suggests that *HOXD12* may function as an oncogene in oligodendroglioma and may drive a more aggressive subtype of oligodendroglioma that occurs disproportionately in patients at an older age [59]. We conclude our work in chapter 8.

Chapter 2

DATA

2.1 Magnetic Resonance Imaging (MRI)

2.1.1 MRI Description

Magnetic resonance imaging (MRI) is a non-invasive medical imaging technique that uses powerful magnetic fields and radio waves to produce detailed images of the body's interior. Standard MRI sequences include pre- (T1) and post-contrast (T1ce) T1-weighted sequences and T2-weighted (T2) and T2 Fluid-Attenuated Inversion Recovery (FLAIR) sequences. The T1 MRI sequence measures the time it takes for the hydrogen nuclei in the body's tissues to return to their resting state after being exposed to a magnetic field and gives a baseline image of the brain's anatomy. The T1 post-contrast MRI sequence is a T1 sequence performed after the injection of a contrast agent into the patient's bloodstream. Contrast agents, such as gadolinium, are substances that can be detected by MRI that enhance the visibility of areas where the brain vasculature has been broken down and are thought to be sites of the tumor's most aggressive activity. The T2 and FLAIR sequences use radiofrequency pulses to invert the magnetization of fluid in the brain, which can suppress the signal from cerebrospinal fluid (CSF) and other fluids. This can make it easier to see certain structures and tumor compartments, such as peritumoral edema.

2.1.2 MRI Datasets

BraTS MRI Dataset

The Multimodal Brain Tumor Segmentation Challenge (BraTS) 2018 dataset provides multi-institutional pre-operative multimodal MRI tumor scans for 542 patients across BraTS' train-

ing, validation, and testing datasets [60, 61, 62, 63, 64]. Each patient is represented by T1, T1ce, T2, and FLAIR MRI sequences. These data are co-registered to the standard MNI anatomical template, interpolated to the same resolution, and skull-stripped. Tumor segmentations are available for the 285 samples in the BraTS training section. Segmentation files provide voxel-wise annotation labels for enhancing tumor, peritumoral edema, background, and necrotic core and non-enhancing tumor compartments. Necrotic core and non-enhancing tumor share a single label. Ground-truth segmentations were manually drawn and approved by neuroradiologists.

TCIA Dataset

MRI data from 46 IDH-wildtype glioblastoma patients with T1, T1ce, T2, and FLAIR MRI sequences were downloaded from The Cancer Imaging Atlas (TCIA) data portal [65]. The Brain Extraction Tool (BET) and FMRIB's Linear Image Registration Tool (FLIRT) from the FMRIB Software Library (FSL) were used to skull-strip and co-register same-subject MRI sequences [66, 67]. All MRI volumes were resampled to $1mm^3$ isotropic space.

2.2 Somatic copy number alterations (SCNA)

2.2.1 SCNA Description

Somatic copy number alterations (SCNAs) are changes in the number of copies of a particular gene or group of genes in the genome. These alterations can be gained or lost, and they can have a significant impact on the function of the affected genes. Copy number alterations can arise from a variety of mechanisms, including chromosomal deletions, duplications, or translocations. They are commonly observed in cancer cells, where they can contribute to the development and progression of the disease.

2.2.2 SCNA Datasets

TCGA Glioma SCNA Dataset

Gene-level glioma somatic copy number alteration (SCNA) data for 1090 patients in the TCGA-GBM and TCGA-LGG projects were downloaded from UCSC Xena. These data were preprocessed from copy number segmentation files using the Genomic Identification of Significant Targets in Cancer 2.0 (GISTIC) algorithm and were aligned to the human genome assembly GRCh37 (hg19) [68]. The thresholded output of the GISTIC algorithm assigns each gene a discrete integer value between -2 and 2 corresponding to gene homozygous deletion (-2), single copy deletion (-1), diploid normal copy (0), low-level amplification (1), or high-level amplification (2). We also downloaded copy number segmentation files for these 1090 patients from the National Cancer Institute's Genomic Data Commons (GDC) Data Portal (https://gdc.cancer.gov/) and processed them with GISTIC ourselves. These applications of GISTIC resulted in one version of the TCGA glioma SCNA data aligned to hg19 (GDC hg19) and the other to human genome assembly GRCh38 (GDC hg38). All TCGA SCNA data is derived from Affymetrix SNP 6.0 arrays.

Capper et al. SCNA Dataset

Illumina 450k methylation IDAT files for 489 non-recurrent diffuse gliomas made available by the authors of Capper et al. [47] were downloaded from the NCBI Gene Expression Omnibus (GEO) under accession number GSE109381. These methylation data were processed into copy number segmentation files using the R packages minfi [69] and conumee [70], and copy number calls were computed from these segmentation files using GISTIC 2.0. This patient cohort consisted of 298 IDH-wildtype glioblastomas, 110 IDH-mutant astrocytomas, and 81 IDH-mutant and 1p/19q-codeleted oligodendrogliomas as determined by criteria in the fourth edition of the WHO classification of CNS tumors.

Jonnson et al. SCNA Dataset

Copy number segmentation files from cBioPortal [71, 72] for 432 primary diffuse glioma patients originally described in Jonsson et al. [73] were downloaded and were processed with GISTIC 2.0. Labels for IDH mutational status and 1p/19q-codeletion status were determined from criteria in the fourth edition of the WHO classification of CNS tumors. This patient cohort consisted of 319 IDH-wildtype gliomas, 63 IDH-mutant astrocytomas, and 50 oligodendrogliomas. Copy number alteration data were derived from targeted sequencing (MSK-IMPACT or FoundationOne) as described in [73].

GLASS SCNA Dataset, 2019 Release

The Synapse API was used to download copy number segmentation files (variants-gatk-seg) for 201 primary diffuse astrocytic glioma in the Glioma Longitudinal AnalySiS (GLASS) dataset (Data Release version 2019-03-28) [74, 75]. We processed these data with GISTIC 2.0. This patient cohort consisted of 143 IDH-wildtype glioblastomas and 58 IDH-mutant astrocytomas determined from available IDH mutation annotations.

REMBRANDT SCNA Dataset

Binary CN4.cnchp files from Affymetrix Human Mapping 50K Hind240 (N=240) and 50K Xba240 SNP arrays (N=192) for 275 samples were downloaded from the REMBRANDT Database (GEO Data Set GSE108475) [76]. Affymetrix Power Tools (February 2021, date last accessed) was used to convert the CN4.cnchp files into text files. Precomputed copy number and loss of heterozygosity analysis results from the CN4 algorithm were extracted from these files, and HmmMedianLog2Ratio values were used to estimate the underlying DNA copy number variation using the Bioconductor package DNAcopy [77]. GISTIC was then applied to calculate gene-level gains and losses. In general, we found that data produced by Hind SNP arrays were cleaner than those generated by Xba SNP arrays and thus we used Hind-derived SCNA data for patients who had both Hind and Xba data (N=157).

NYU SCNA Dataset

Copy number segmentation files for 68 IDH-wildtype glioblastoma patients were obtained from an institutional dataset at New York University (NYU). Thresholded gene-level copy number calls were computed using the GISTIC 2.0 algorithm aligned to hg19. Age (N=68) and MGMT promoter methylation status (N=63) were available.

2.3 Gene Expression

2.3.1 Gene Expression Description

Gene expression is the process by which the information encoded in a gene is used to produce a functional gene product, such as a RNA molecule. This process involves transcribing the DNA code for a gene into a complementary RNA molecule. Gene expression is essential for cells to function properly, as proteins translated from RNA perform a wide range of functions in the body, including catalyzing chemical reactions, providing structural support, and regulating cellular processes. The level of gene expression can be regulated by a variety of mechanisms, such as DNA methylation and the binding of regulatory proteins. Gene expression is typically quantified using RNA sequencing (RNAseq) or single-cell RNA sequencing (scRNAseq), which measures gene expression at the cellular level.

2.3.2 Gene Expression Datasets

TCGA Bulk RNAseq Dataset

RNA-seq gene counts for TCGA adult diffuse glioma were downloaded from the National Cancer Institute's Genomic Data Commons (GDC) Data Portal (`https://gdc.cancer.gov/`). TCGA transcript per million data (TPM) were computed from the recount2 dataset [78] for 659 adult diffuse glioma, as specified by Arora et al. [79] to be consistent with the available Chinese Glioma Genome Atlas (CGGA) TPM data. These tumors include 234 IDH-wildtype gliomas, 256 IDH-mutant astrocytomas, and 169 oligodendrogliomas.

CGGA Bulk RNAseq Dataset

Raw RNA gene counts for 190 Chinese Glioma Genome Atlas (CGGA) [80] primary adult diffuse glioma tumors were downloaded from recount2 [78] and used to compute TPM values [79]. These tumors include 84 IDH-wildtype gliomas, 57 IDH-mutant astrocytomas, and 48 oligodendrogliomas. Unsupervised dimension reduction [81] of CGGA and TCGA TPM data identified apparent outliers, which were excluded from our gene expression analysis (Supplemental Figure A.19).

GLASS Bulk RNAseq Dataset

The Synapse API was used to download gene expression TPM data for 147 GLASS adult-type diffuse glioma patients (Data Release version 2021-11-15) [75, 74]. This cohort consisted of 116 primary IDH-wildtype glioblastomas, 21 primary IDH-mutant astrocytomas, and 10 primary oligodendrogliomas. All GLASS samples that appear in the TCGA were excluded.

2.4 DNA Methylation

2.4.1 DNA Methylation Description

DNA methylation is a process by which methyl groups are added to DNA. These epigenetic DNA modifications can change the activity of a gene without changing its DNA sequence. Methylation typically occurs in the promoter region or body of genes, where it can play a role in regulating gene expression.

2.4.2 DNA Methylation Datasets

TCGA DNA Methylation Dataset

Precomputed TCGA DNA methylation beta values (Illumina 450K) from 658 adult diffuse gliomas were downloaded from University of California Santa Cruz (UCSC) Xena (`https://`

xena.ucsc.edu/), including 248 IDH-wildtype glioblatoma, 197 IDH-mutant astrocytomas, and 171 primary oligodendrogliomas [82].

Capper et al. DNA Methylation Dataset

Illumina 450K platform DNA methylation IDAT files for 3905 CNS tumors made available by the authors of Capper et al. [47] were downloaded from the NCBI Gene Expression Omnibus (GEO) under accession number GSE109381. Beta values were determined using the R package minfi [69]. By the WHO 2016 classification of CNS tumors, 768 were IDH-wildtype glioma, 250 were IDH-mutant astrocytomas, and 170 were oligodendrogliomas. By the classification of the version 11 Capper et al. random forest classifier or if the reported final pathological diagnosis (after integration with the random forest classifier result) 529 were IDH-wildtype glioblastomas, 252 were IDH-mutant astrocytomas, and 162 were oligodendrogliomas.

GLASS Bulk DNA Methylation Dataset

The Synapse API was used to download DNA methylation beta values from the merged Illumina 450K and EPIC platforms (beta.merged) for 46 IDH-wildtype glioblastomas, 46 IDH-mutant astrocytomas, and 11 oligodendrogliomas (Data Release version 2021-11-15). All GLASS samples that appear in the TCGA were excluded.

2.5 SNP

2.5.1 SNP Description

Single nucleotide polymorphisms (SNPs) are the most common type of genetic variation; they occur when a single nucleotide (A, C, G, or T) in the DNA sequence is altered. These variations are relatively common in the human genome, and they can have a variety of effects, depending on where they occur and how they alter the sequence of nucleotides. Some SNPs

have no effect, while others can alter the function of a gene leading to the activation of an oncogene or the inactivation of a tumor suppressor.

2.5.2 SNP Datasets

TCGA SNP Dataset

Somatic mutation calls for The Cancer Genome Atlas (TCGA) glioblastomas and lower-grade astrocytic and oligodendroglial tumors (N=812) computed by the Multi-Center Mutation Calling in Multiple Cancers (MC3) project [83] were downloaded from University of California Santa Cruz (UCSC) Xena (https://xena.ucsc.edu/) [82].

2.6 Histopathologic Features

2.6.1 Description

Common histopathologic features described by neuropathologists in pathology reports include Ki-67 proliferation index, necrosis, microvascular proliferation, and mitotic figures. The Ki-67 proliferation index is a measure of cancer cells that are actively dividing and a marker of tumor aggressiveness. Necrosis refers to areas of dead cancer cells within the tumor. Microvascular proliferation refers to the agressive growth of new blood vessels within the tumor. Mitotic figures are cells in the process of dividing.

2.6.2 Datasets

TCGA Histopathologic Features

TCGA pathology reports for 156 oligodendroglioma patients were downloaded from the GDC Legacy Archive. Values for the Ki-67 proliferation index, presence of necrosis, presence of microvascular proliferation, and presence of mitotic figures were recorded. When the Ki-67 index was given as a range, we reported the upper bound. Ki-67 index values were considered high ($\geq$10%) or low ($<$10%) based on published reports [84]. We considered $\geq$6 per 10 high

power fields (HPF) high and <6 per 10 HPF low per WHO guidelines.

Chapter 3

3D-ESPNET WITH PYRAMIDAL REFINEMENT FOR VOLUMETRIC BRAIN TUMOR IMAGE SEGMENTATION

Abstract

Automatic quantitative analysis of structural magnetic resonance imaging (MRI) of brain tumors is critical to the clinical care of glioma patients, and for the future of advanced MRI research. In particular, automatic brain tumor segmentation can provide volumes of interest (VOIs) to scale the analysis of advanced MRI imaging modalities such as perfusion-weighted imaging (PWI), diffusion-weighted imaging (DTI), and MR spectroscopy (MRS), which is currently hindered by the prohibitive cost and time of manual segmentations. However, automatic brain tumor segmentation is complicated by the high heterogeneity and dimensionality of MR data, and the relatively small size of available datasets. This paper extends ESPNet, a fast and efficient network designed for vanilla 2D semantic segmentation, to challenging 3D data in the medical imaging domain. Even without substantive pre- and post-processing, our model achieves respectable brain tumor segmentation results, while learning only 3.8 million parameters. 3D-ESPNet achieves dice scores of 0.850, 0.665, and 0.782 on whole tumor, enhancing tumor, and tumor core classes on the test set of the 2018 BraTS challenge. Our source code is open-source and available at `https://github.com/sacmehta/3D-ESPNet`.

3.1 Introduction

Glioma is the most common primary malignant adult brain tumor. Due to glioma's highly heterogeneous appearance, extent, and shape, segmentation of brain tumors in MRI volumes is one of the most challenging tasks in neuroradiology [85]. This is compounded by the sparsity of data and the heterogeneity incurred by differing scanner models and manufacturers,

imaging sites, variation in clinical standards and protocols, and the noise introduced by the movement of patients' heads during scans. At every clinical visit, glioma patients generally receive standard-of-care FLAIR, post-contrast T1-weighted (T1ce), T2, and T1 MRI sequences, each of which is described by a distinct volume. These sequences give distinct and complementary information about the tumor's extent and composition.

Automated brain tumor segmentation also ranks among the most difficult problems in medical image analysis. The notion that massive amounts of data are required to train deep networks is widely held. Not only are MRI scans scarce, they are high dimensional (e.g., $240 \times 240 \times 155 \times 4$) and contain high class imbalances (e.g., $\geq 95\%$ background class). Thus, naive models are predisposed to exhibit extreme background bias.

In similar biomedical domains, patchwise approaches have helped address problems of data shortages and dimensionality. Ciresan et al. proposed a sliding-window method to segment electron microscopic images of the brain, which both localized the problem and exaggerated the dataset [86, 87]. Ronneberger et al.'s 2D encoder-decoder network, U-Net, outperformed Ciresan's method [86]. U-Net is a fully convolutional network (FCN) where the traditional pooling operations in the contracting (encoding) path are mirrored by upsampling operations in the symmetric expanding (decoding) path. Skip connections are passed from encoding blocks on the contracting path to same-level decoding blocks in the expanding path.

While some success has been reached using 2D FCNs, like U-Net, these models ignore crucial 3D spatial context, which is undesirable given that most clinical imaging data are volumetric. However, even among 3D FCNs such as DeepMedic, a previous winner of the BraTS competition, fine spatial information is discarded in pooling [88]. This motivates our interest in U-Net's skip connections and, in particular, the architecture of Milletari et al.'s 3D extension of U-Net, V-Net. V-Net benchmarked well on the "PROMISE2012" challenge, where it gave impressive segmentations of MRI prostate scans after training on only 50 examples [89].

ESPNet is a faster, more efficient take on U-Net's encoder-decoder architecture [90]. In

this paper, we seek to extend and benchmark ESPNet on 3D medical imaging data.

We outline our paper as follows. Section 2 describes our network architecture. We report our methods in Section 3. Experimental results are given in Section 4. Finally, we close with a discussion of limitations and future directions for our work in Section 5.

3.2 Methods

3.2.1 Network Architecture

Our network is an end-to-end system consisting of 3D-ESPNet followed by pyramidal refinement, as shown in Figure 3.1. We describe the main building block of our architecture, the ESP module, and, later, 3D-ESPNet's segmentation architecture and pyramidal refinement.

3.2.2 ESP Module

The Efficient Spatial Pyramid (ESP) module, shown in Figure 3.2, is an efficient convolutional module proposed in [90]. The module is based on the RSTM (Reduce-Split-Transform-Merge) strategy and allows the aggregation of the information from a large effective receptive field while learning fewer parameters. We extend the ESP block by replacing its spatial 2D convolutions with volumetric 3D convolutions.

3.2.3 3D-ESPNet Structure

3D-ESPNet is an encoder-decoder network that extends U-Net [86]. The primary distinction between 3D-ESPNet and U-Net is that 3D-ESPNet employs efficient convolutional blocks for aggregating features instead of stacking convolution layers (with or without residual connections) after the first layer.

In the encoder stage, the network learns feature representations by performing convolutional and downsampling operations. The encoder downsamples once with a strided convolutional layer and three subsequent times with strided ESP modules. In downsampling ESP modules, we use convolutions with $n_i \times n_i \times n_i$ sized kernels and stride of two, for

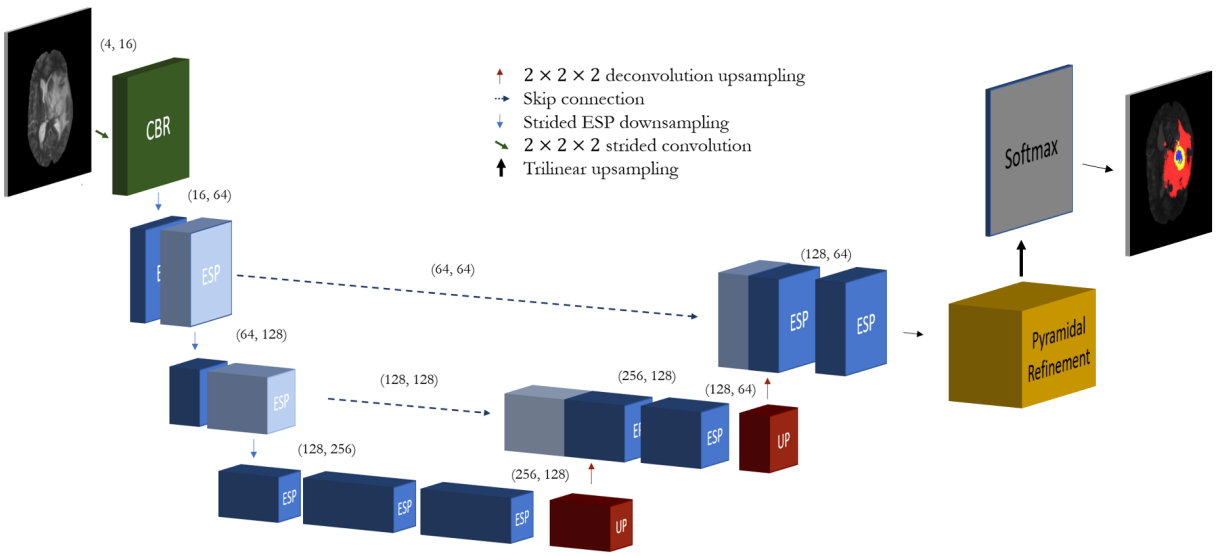


Figure 3.1: 3D-ESPNet with pyramidal refinement. 3D-ESPNet's encoder is shown on the left; the decoder is shown on the right with pyramidal refinement. Parentheses give the channel dimensions of incoming and outgoing feature maps. The CBR block consists of a convolutional block followed by batch normalization and ReLU. Light-blue feature maps in the decoder indicate concatenation by long-range, skip connections. Light-blue feature maps in the encoder indicate strided ESP models for downsampling. Arrows are defined in the legend.

$i \in \{1, \ldots, K\}$, as shown in Figure 3.2. The combination of varying receptive fields allows 3D-ESPNet to learn feature representations at multiple scales.

In the decoder stage, we share the feature maps in the encoder with same-level feature maps in the decoder via skip-connection concatenation. Skip-connections allow fine details lost in downsampling in the encoder to be recovered in the decoder, which gives the segmentation maps a granularity that simple interpolation cannot achieve. The decoder uses $3 \times 3 \times 3$ deconvolution kernels to upsample the encoder output once, followed by a trilinear upsampling layer to return to the resolution at the network's second level. The feature maps of the final ESP module in the decoder are passed into the pyramidal refinement module.

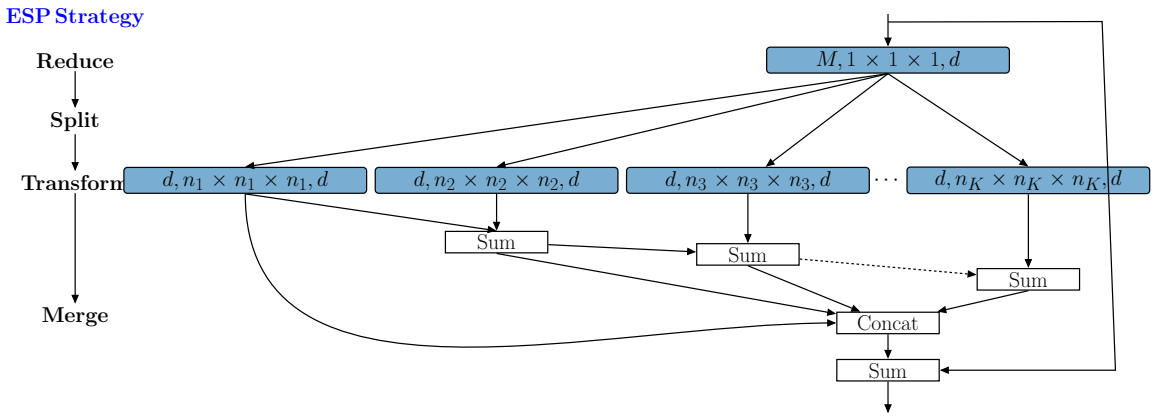


Figure 3.2: The Efficient Spatial Pyramid (ESP) module. The blocks in blue represent 3D convolutional layers and are denoted as (# input channels, effective receptive field, # output channels). The ESP module takes an input feature map with M channels and produces an output feature map with N channels, where $d = \frac{N}{K}$ and K represent the number of parallel branches.

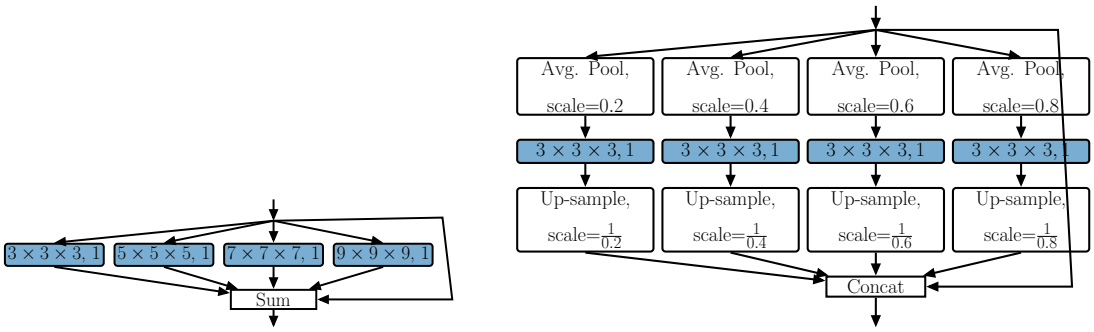
The block diagram of 3D-ESPNet is shown in Figure 3.1.

Pyramidal refinement: Pyramid-based approaches sub-sample either the feature maps or the convolutional kernel to learn global contextual information. Inspired by the success of such approaches for segmenting complex 2D scenes, we extend these modules for volumetric data. We call this module pyramidal refinement. Our module combines both feature map-based and convolutional kernel-based pooling methods in a novel fashion.

Pyramidal refinement, shown in Figure 3.4, consists of three layers:

- *Projection Layer:* This is a standard $3 \times 3 \times 3$ convolutional layer followed by batch normalization and a ReLU function that projects the feature maps from the previous ESP block to C-dimensional space, where C is the number of classes.

- *Spatial Pyramid pooling (SPP) block*: The input feature maps to this block are low



(a) SPP block. (b) PSP block.

Figure 3.3: Pooling modules used in a pyramidal refinement block. Here, a convolutional layer is represented as (kernel size, dilation rate).

dimensional ($C = 4$). We sub-sample them using convolutional kernels of different sizes and merge their output using sum operations. This is similar to the ASPP block except that we do not use dilated convolutions [91].

- *PSP block:* A PSP block, sketched in Figure 3.3b, is based on the principle of *split-pool-transform-upsample* [92]. *Split:* A PSP block distributes the input feature maps across four parallel branches. *Pool:* Each branch downsamples the feature maps using a different pooling rate. *Transform:* The downsampled feature maps are transformed using point-wise convolutions. *Upsample:* The transformed feature maps are upsampled to the same resolution as the input feature maps using bilinear interpolation. *Merge:* The upsampled feature maps are concatenated with the input feature maps to produce the output feature maps.

Pyramidal refinement is followed by a classification layer. This final layer pools the feature maps using another SPP block and then upsamples by a factor of two using trilinear interpolation. Two convolutional layers are stacked on top of the upsampled feature maps before a softmax.

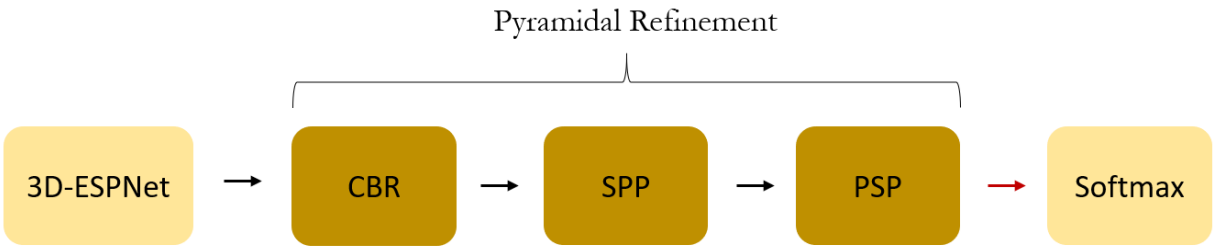


Figure 3.4: Pyramidal-refinement. After the second upsampling operation in the 3D-ESPNet decoder, the feature maps are passed through a CBR block, a spatial pyramid pooling block (SPP), and a pyramid pooling module (PSP) at 1/4 resolution. We then upsample to input resolution using trilinear interpolation and compress and pass the feature maps through a softmax to obtain a prediction.

3.2.4 Data

We train on the Multimodal Brain Tumor Segmentation Challenge (BraTS) 2018 training set, which provides 285 multi-institutional pre-operative multimodal MRI tumor scans, each consisting of T1, post-contrast T1-weighted (T1ce), T2, and FLAIR volumes [60, 61, 62, 63, 64]. Each case is annotated with the following voxel labels: enhancing tumor, peritumoral edema, background, and necrotic core and non-enhancing tumor. Necrotic core and non-enhancing tumor share a single label. These data are co-registered to the standard MNI anatomical template, interpolated to the same resolution, and skull-stripped. Ground-truth segmentations are manually drawn and approved by neuroradiologists.

Preprocessing

We used minimal preprocessing. We performed min-max normalization. We also cropped each volume to remove any padding around the brain common in every modality; this allowed us to double our batch size to four, which stabilized training.

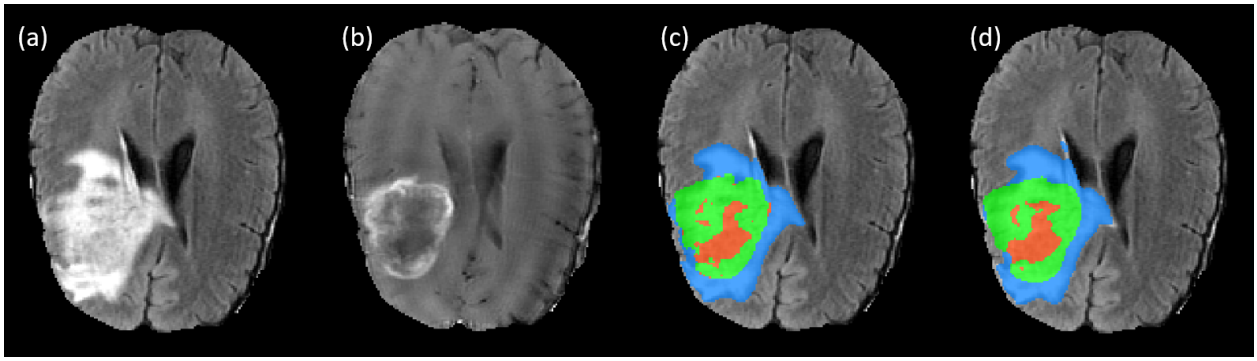


Figure 3.5: A butterfly high-grade glioma. (a) FLAIR sequence; (b) T1ce sequence; (c) network prediction; (d) ground truth segmentation.

3.2.5 Training

To tune our model's hyperparameters, we randomly partitioned our dataset into a training set and a validation set using an 80:20 split (228:57). We selected the hyperparameters that maximized the mean intersection over union (mIOU) on the 57 withheld volumes in the validation set. We used mean intersection over union (mIOU) for our loss function instead of cross entropy for empirical reasons as we and others have observed [93]. We weight our mIOU loss to address the severe class imbalance. We used data augmentation heavily including scaling and random flips.

We implemented our model in PyTorch. We trained at full resolution on all modalities on an NVIDIA Titan X using a batch size of four. We trained for 300 epochs. Training took less than five hours; test time evaluation takes less than twenty seconds. We found that the optimizer Adam outperformed SGD with momentum [94]. We experimented with learning rate decay and settled on a learning rate of $10e^{-4}$, which we decreased to $10e^{-5}$ after 200 epochs. Code for this adaptation of ESPNet is available at `https://github.com/sacmehta/3D-ESPNet`.

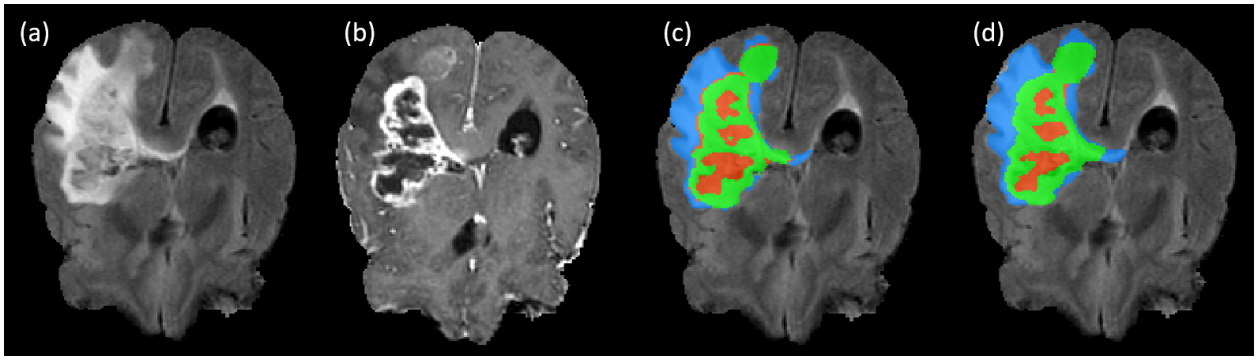


Figure 3.6: A second butterfly high-grade glioma. (a) FLAIR sequence; (b) T1ce sequence; (c) network prediction; (d) ground truth segmentation.

3.3 Results

Results on the BraTS 2018 online test and validation sets are shown in Table 1. Visual inspection reveals out model's flexible performance on difficult cases such as gliomas that cross the corpus callosum–so-called butterfly gliomas–shown in Figures 3.5 and 3.6. However, our method lacks some of the granularity present in the ground truth segmentation. It is clear in the examples provided that our network's predictions are too smooth, especially in Figure 3.5, where the predicted non-enhancing and necrotic class is the correct size and in the correct position, but the segmentation does not follow the sharp contours of the gyri outlined in the ground truth. In Figure 3.6, we notice that our network tends not to predict necrotic or non-enhancing tumor outside of the tumor-enhancing ring. However, our model is able to handle gaping holes inside tumors filled with cerebrospinal fluid (CSF) just as a resection cavity would appear. This is shown in Figure 3.8. These cavities differ from a typical necrotic core on the T2 sequences of a tumor as CSF shows extreme hyperintensity. This robustness is crucial for segmenting post-operative scans which can contain large resection cavities.

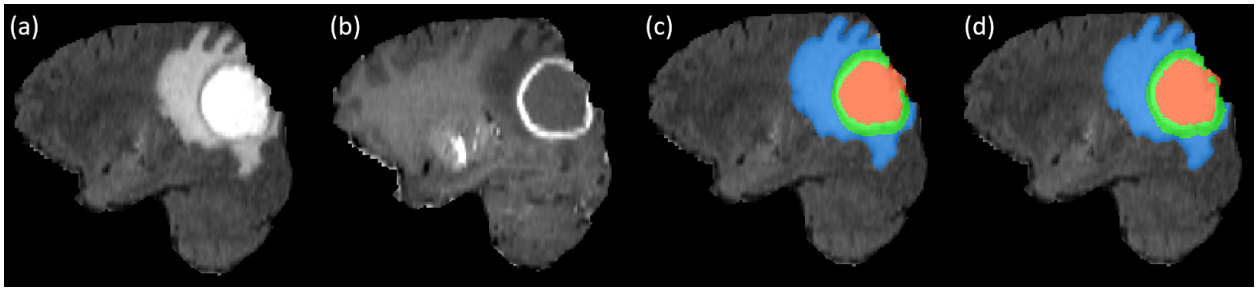


Table 3.1: Results obtained on BraTS 2018 online test set are shown in bold. Results obtained on BraTS 2018 online validation set are shown in parentheses. Sensitivity and specificity results were not given for the online test set.

3D-ESPNet	**Dice Score**		**Sensitivity**		**Specificity**		**Hausdorff95**	
Whole Tumor	**0.850**	(0.883)	**-**	(0.934)	**-**	(0.990)	**9.598**	(5.461)
Enhancing Tumor	**0.665**	(0.737)	**-**	(0.831)	**-**	(0.997)	**5.497**	(5.295)
Tumor Core	**0.782**	(0.814)	**-**	(0.821)	**-**	(0.997)	**8.668**	(7.850)

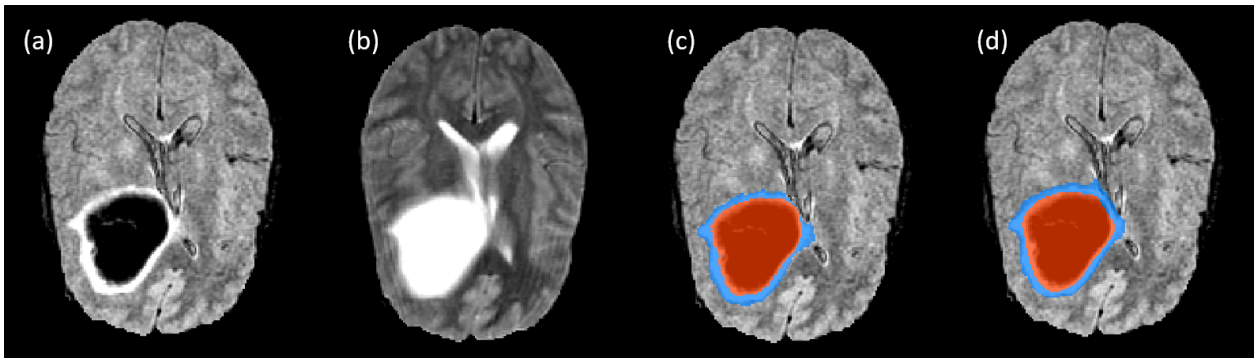


Figure 3.8: Low-grade glioma showing bright CSF fluid in ventricles and tumor cavity on the T2 sequence. (a) FLAIR sequences; (b) T2 sequence; (c) network prediction; (d) ground truth segmentation.

3.4 Discussion

We propose a fast and efficient network for semantic brain tumor segmentation. 3D-ESPNet with pyramidal refinement achieves a respectable 0.850 dice score for whole tumor segmentation on the 2018 BraTS online test set without substantial pre- or post-processing, while learning only 3.8 million parameters.

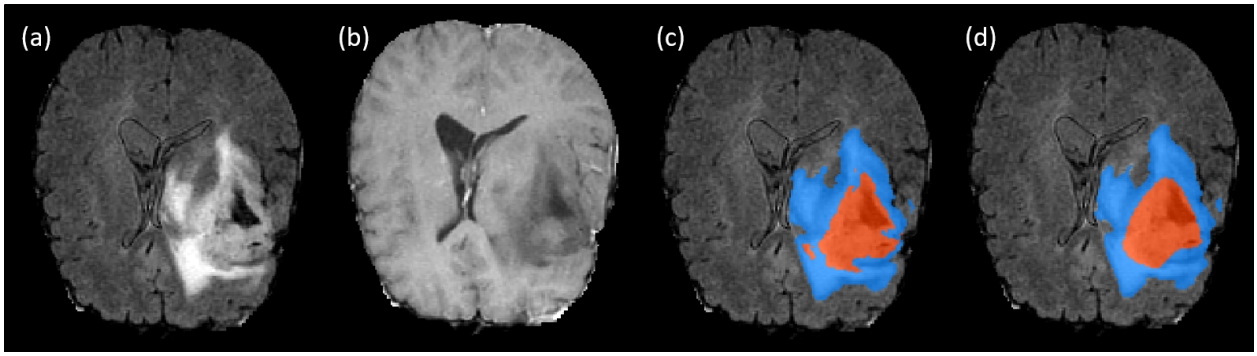


Figure 3.9: Low-grade glioma. (a) FLAIR sequence; (b) T1 sequence; (c) network prediction; (d) ground truth segmentation.

Brain tumor segmentation has its place in clinic, though neuroradiologists and neuro-oncologists usually limit its use to quantifying volumetric changes in tissue types (edema, enhancing tissue, non-enhancing or necrotic tissue) between patient visits for evaluating tumor progression [85]. However, tumor segmentation is essential to the analysis of advanced MRI (DWI, DTI, MRSI). Because such segmentation is usually done manually, segmentation time and cost prevent advanced MRI studies from being done at scale. Automatic brain tumor segmentation will allow such advanced imaging studies to be done on massive datasets and, therefore, avail themselves of strong ML analysis and more definitive conclusions.

We plan to add pre- and post-processing techniques to our model. Histogram equalization and N4BiasFieldCorrection might better prepare the training data, and adding a conditional random field after the classifier may help eliminate spurious tumor predictions. We achieved a dice score of 0.850 on the whole tumor class, but work remains to be done on the individual classes. Better hyperparameter tuning and non-linear data augmentation may also improve our performance.

Chapter 4

RADIOGENOMIC MODELING PREDICTS SURVIVAL-ASSOCIATED PROGNOSTIC GROUPS IN GLIOBLASTOMA

Abstract

Background: Combined whole-exome sequencing (WES) and somatic copy number alteration (SCNA) information can separate IDH1/2-wildtype glioblastoma into two prognostic molecular subtypes, which cannot be distinguished by epigenetic or clinical features. The potential for radiographic features to discriminate between these molecular subtypes had yet to be established. *Methods:* Radiologic features (n=35,340) were extracted from 46 multi-sequence, pre-operative magnetic resonance imaging (MRI) scans of IDH1/2-wildtype glioblastoma patients from The Cancer Imaging Archive, all of whom have corresponding WES/SCNA data. We developed a novel feature selection method that leverages the structure of extracted MRI features to mitigate the dimensionality challenge posed by the disparity between the large number of features and the limited patients in our cohort. Six traditional machine learning classifiers were trained to distinguish molecular subtypes using our feature selection method, which was compared to LASSO feature selection, recursive feature elimination, and variance thresholding. *Results:* We were able to classify glioblastomas into two prognostic subgroups with a cross-validated area under the curve score of 0.80 ($\pm$0.03) using ridge logistic regression on the 15-dimensional principle component analysis (PCA) embedding of the features selected by our novel feature selection method. An interrogation of the selected features suggested that features describing contours in the T2 signal abnormality region on the T2-weighted FLAIR MRI sequence may best distinguish these two groups from one another. *Conclusions:* We successfully trained a machine learning model that allows for relevant

targeted feature extraction from standard MRI to accurately predict molecularly-defined risk-stratifying IDH1/2-wildtype glioblastoma patient groups.

4.1 Introduction

Glioblastoma is a highly aggressive disease that is largely characterized by somatic copy number alterations (SCNAs), which are changes in chromosome structure resulting in gains or losses of either regions of chromosomes (e.g., EGFR amplification, CDKN2A/B deletion, loss of chromosomal region 9p) or whole chromosomes (e.g., gain of chromosome 7, loss of chromosome 10) [38, 45, 15, 46, 95]. Certain SCNA signatures in glioblastoma have demonstrated prognostic utility with implications for risk-stratification beyond conventional histological grading [15, 96, 97, 98, 99]. Using dimensionality reduction mapping of combined whole exome single nucleotide mutations and exome-wide SCNAs, diffuse gliomas from The Cancer Genome Atlas (TCGA) display regional mapping clusters [15]. Projection of a second independent cohort of paired initial and recurrent glioblastomas onto the TCGA reference map additionally indicates that there are genomic-based regions where patients have glioblastomas that are deemed unresectable at recurrence (Group 1) (Figure 4.1A). Furthermore, the isocitrate dehydrogenase (IDH) 1 and 2 wildtype glioblastoma patients in this Group 1 have worse overall survival than those IDH-wildtype glioblastoma patients represented in the region overlapping with the paired cohort (Group 2) (Figure 4.1B). Group 1 tends to be relatively genomically stable with respect to copy number alterations beyond gain of chromosome 7, loss of chromosome 10, and loss of the chromosomal 9p region. We have previously shown that assignment to one of these two genomically-determined groups (Figure 4.1A, C) is not predicted by clinical factors (age, sex, Karnofsky Performance Status) or epigenetic signatures (including genome-wide methylation and gene expression) [15]. The role of radiology in distinguishing these two groups is unknown. Given the implications of group membership to patient outcome, along with the relative abundance of pre-operative magnetic resonance imaging (MRI) and the lack of routine whole exome sequencing of glioblastoma samples, we sought to determine whether radiographic features can differentiate Group 1 from Group

2. To do this, we turned to a cohort of IDH-wildtype glioblastomas with multi-sequence, pre-operative MRI in The Cancer Imaging Archive (TCIA) and used radiogenomic methods to train machine learning classifiers to discriminate between these two groups.

Radiogenomics is an evolving field in medical imaging that employs supervised and unsupervised learning to relate quantitative imaging features to the underlying genomic characteristics of the imaged tissue [23]. Radiogenomic pipelines typically consist of image acquisition, image normalization, feature extraction, and prediction using either coupled feature selection and machine learning models or end-to-end prediction using deep convolutional neural networks [24, 25]. In neuro-oncology, radiogenomic approaches have been used to predict IDH mutations [25, 26, 27, 28, 29], 1p/19q-codeletion [29, 30], O6-methylguanine-DNA methyltransferase (MGMT) promoter methylation [30, 31], several gene-level SCNAs [32], and tumor treatment response [33]. Though powerful, radiogenomic methods in neuro-oncology are prone to overfit classification tasks because MRI data is high-dimensional, relatively scarce, extremely heterogenous, and inherently noisy due to variations in MRI scanner manufacturer, sequence protocol, and patient movement. This is especially true for complex machine learning models, such as deep neural networks, which excel in certain glioma imaging tasks, such as tumor segmentation, but struggle with nuanced classification and regression tasks, such as survival prediction [34, 36]. Even traditional machine learning models can be overwhelmed by the size of the feature spaces extracted by open-source radiogenomic software ($N > 10,000$) [100, 101]. To reduce dimensionality, standard feature selection methods such as LASSO feature selection, recursive feature elimination, and variance thresholding are commonly employed. However, these methods can also overfit classification tasks in small studies where predictions must be made on large, noisy data. To address this, we propose a novel feature selection method that allows simple machine learning models to accurately predict Group 1 and Group 2 and gives insight into what radiographic features discriminate these two molecularly-defined groups from one another.

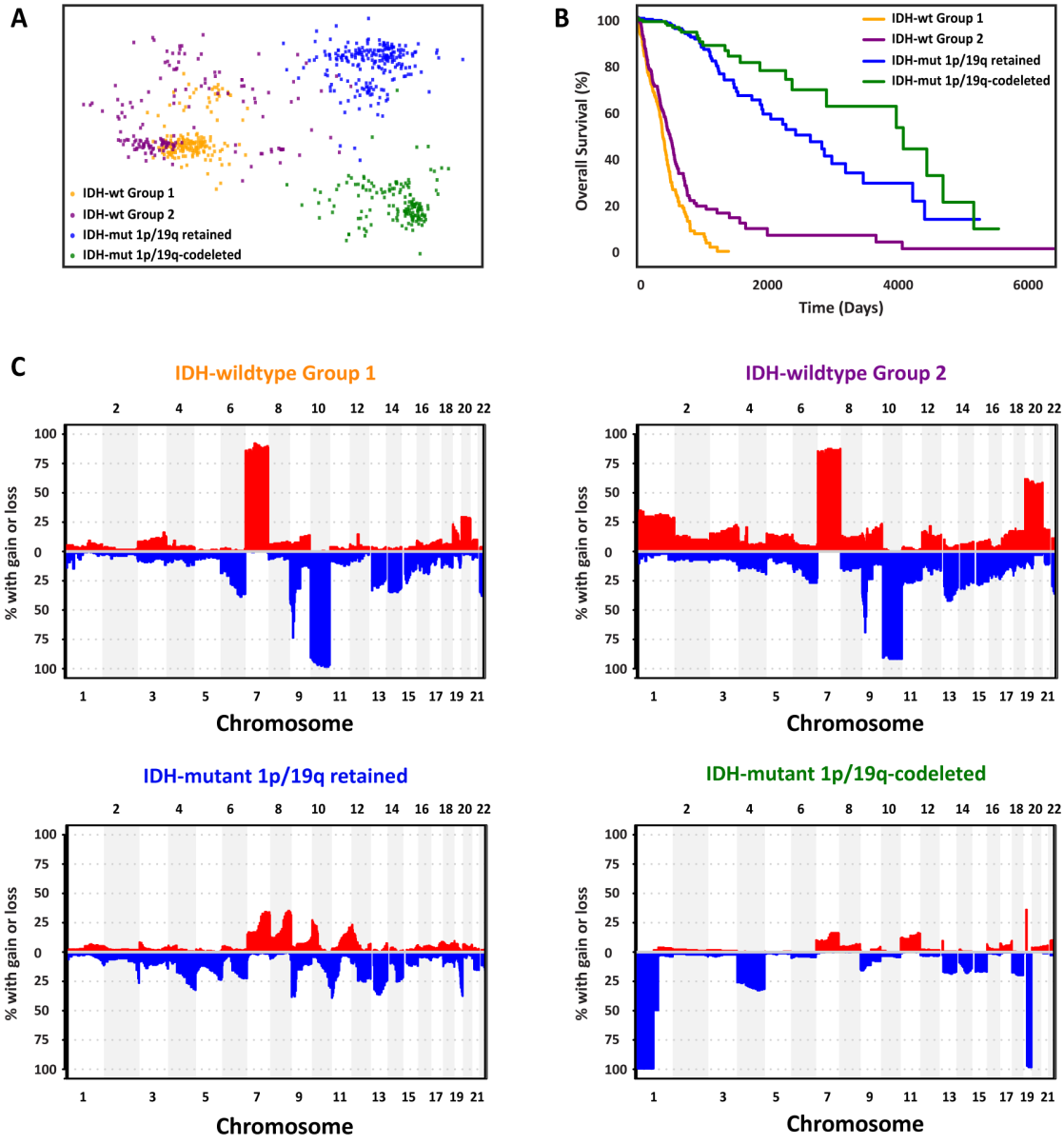


Figure 4.1: Characteristics of genomically-defined prognostic adult-type diffuse glioma subtypes. (A) Multidimensional scaling plot of IDH-wildtype and IDH-mutant diffuse gliomas from the TCGA. Each point represents a single patient and is determined by the combination of whole exome sequencing and somatic copy number alterations. The IDH-wildtype gliomas are divided into two previously determined prognostic subtypes (Group 1 and Group 2). (B) Kaplan-Meier curves for IDH-wildtype (Groups 1 and 2) and IDH-mutant gliomas. (C) Copy number frequency plots of diffuse glioma subtypes.

4.2 Data Sets

TCGA Glioma Datasets: Genomic whole exome sequencing and somatic copy number alteration data for The Cancer Genome Atlas (TCGA) glioblastomas, as well as lower-grade astrocytic and oligodendroglial tumors, were obtained from the University of California Santa Cruz cancer browser (`https://genome-cancer.ucsc.edu/`) as previously described [15, 46]. From the whole exome sequencing data, we incorporated non-synonymous nucleotide point mutations into our analysis as previously described [15, 46, 95]. For copy number alterations, SNP-array derived GISTIC 2.0 scores were used for analysis [15, 46, 95]. Dimensionality reduction and visualization of the exome-wide point mutations combined with copy number alterations was performed for each patient using multidimensional scaling (MDS) as previously described [15, 46, 95]. IDH-wildtype diffuse astrocytic gliomas are divided into two prognostic groups, which we label Group 1 and Group 2, based upon a previously published MDS landscape and mapping of an independent cohort of paired initial and recurrent glioblastomas [15]. Of these IDH-wildtype gliomas, we considered only the subset of the World Health Organization (WHO) grade 4 glioblastomas.

TCIA Magnetic Resonance Imaging: Multi-sequence MRI volumes for patients with IDH-wildtype glioblastoma were downloaded from The Cancer Imaging Atlas (TCIA) data portal (`https://public.cancerimagingarchive.net/nbia-search/`) [65]. Of these patients, 46 met the inclusion criteria requiring available SCNA and WES data from the TCGA and usable pre-operative pre- (T1) and post-contrast (T1ce) T1-weighted sequences along with T2-weighted (T2) and T2 Fluid-Attenuated Inversion Recovery (FLAIR) sequences [102]. Group 1 consisted of 25 patients (median age 59 years; 10 female, 15 male); Group 2 consisted of 21 patients (median age 63 years; 9 female, 12 male). The Brain Extraction Tool (BET) and FMRIB's Linear Image Registration Tool (FLIRT) from the FMRIB Software Library (FSL) were used to skull-strip and co-register same-subject MRI sequences [66, 67]. All MRI volumes were resampled to 1mm3 isotropic space. Each tumor was automatically segmented into subcompartments using a publicly available segmentation model

[57] pre-trained on the 2018 Multimodal Brain Tumor Segmentation Challenge (BraTS) MRI dataset [34, 103, 102]. This model placed each tumor voxel into one of three tumor regions used in the BraTS challenges' segmentation scheme: the contrast-enhancing tumor compartment, the non-enhancing tumor and necrotic tissue, and the peritumoral edema tumor region. We modified our predicted segmentation maps to conform to a different segmentation scheme we regard as more intuitive. We maintained the same contrast-enhancing tumor compartment definition as the BraTS challenges, but we split the non-enhancing tumor and necrotic tissue regions. We considered necrotic tissue as a tumor region on its own, and we merged the non-enhancing tumor region with the BraTS challenges' peritumoral edema tumor region to form a tumor region referred to as the T2 abnormality. In our view, non-enhancing tumor tissue is challenging to distinguish from peritumoral edema, which must be done in the BraTS segmentation scheme. In our scheme, however, bright contrast-enhancing tumor and dark necrotic tissue are easily identifiable on the T1ce sequence; all other tissue appearing abnormal on the T2 or FLAIR sequence is, by definition, the T2 abnormality tumor region. In addition to the enhancing tumor region, necrotic tissue, and the T2 abnormality, we also considered the tumor core, formed by merging the enhancing and necrotic regions, and the entire tumor as additional tumor regions. All segmentation masks were examined and manually corrected using ITK-SNAP to the satisfaction of at least one experienced neuro-radiologist at The University of Washington Medical Center [104]. Finally, N4 bias field correction and min-max normalization were applied to the MRI data [105].

4.3 Methods

MRI Feature Extraction: To build a prediction model based upon radiographic signatures, we first represented each patient by a set of quantitative MRI features. To do this, we extracted semantically interpretable histogram and texture features from augmented volumes of interest (VOIs) formed by the application of an image transformation to a tumor region on each MRI sequence. We formed 380 VOIs by independently applying 19 image transformations to the 5 tumor regions in each of the 4 MRI sequences. A total of 35,340

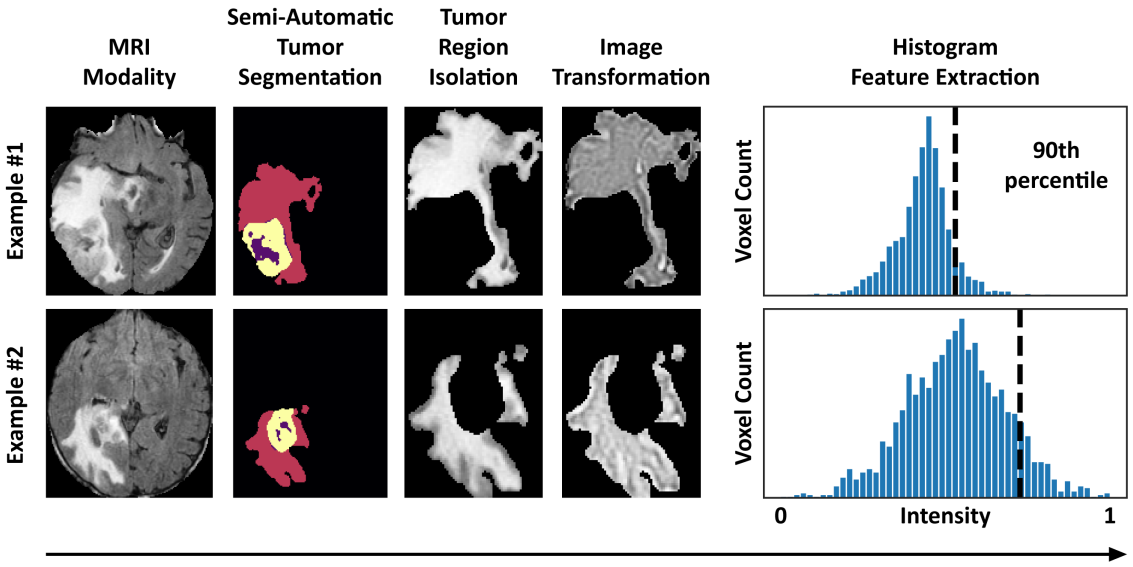


Figure 4.2: Diagram of radiomic feature extraction from MRI. Feature extraction begins by categorizing the MRI sequence (FLAIR imaging shown in these examples). The sequence images then undergo semi-automatic tumor segmentation followed by tumor region isolation. The isolated tumor region then undergoes an image transformation (Laplacian of Gaussian transformation shown in these examples). Here, a histogram feature is derived from the intensity values of the tumor regions and the nth percentile value. This defines an extracted feature.

radiographic histogram/texture features were extracted per patient using 93 histogram and texture features. Our feature extraction pipeline is illustrated in Figure 4.2. Texture features consist of formulas that describe the distribution of values on second-order texture matrices such as the grey level co-occurrence matrix (GLCM) [106]. Examples of image transformations include local binary patterns (LBP), Laplacian of Gaussian (LoG) filters, and Haar wavelet transformations [107, 108]. For later comparison, 60 shape features that describe the VOIs, such as tumor region volume and surface area, and 16 Visually AcceSAble Rembrandt Images (VASARI) MRI features [109], which are manually extracted features designed to standardize visual descriptions of gliomas, were also extracted to ascertain whether visually

perceptible tumor characteristics can distinguish Group 1 from Group 2.

MRI Feature Selection: Following the feature extraction process, we sought to identify the subset of features that are most important to the prediction task. We developed a feature selection method that prioritized identifying characteristics of radiographic features—including MRI sequence, tumor region, image transformation, and histogram/texture formulas—rather than specific radiographic features. We refer to these characteristics as feature components. Because VOIs are each defined by a combination of feature components, our strategy amounts to selecting a set of VOIs and a set of texture/feature formulas that can extract numeric radiogenomic features from these VOIs. The development of our method was guided by our expectation that discriminative radiogenomic features were contained in only a few feature component-defined VOIs and captured by only a few histogram/texture formulas. To select a subset of feature components, we trained LASSO models on random subsets of the training data and used the features they selected to aggregate a bag of repeatedly selected radiographic features. We then broke down the radiographic features in this aggregated bag into their feature components and selected the MRI sequences, tumor regions, image transformations, and histogram/texture formulas that appeared most frequently. Finally, from this set of selected feature components, we identified the set of radiographic features defined by the selected MRI sequences, tumor regions, image transformations, and histogram/texture formulas we selected. We trained our ultimate classification models on a 15-dimensional embedding of this feature set produced by PCA. Our feature selection process is summarized in Figure 4.3A.

Formally, in the first stage of our feature selection method, we aggregated a bag B of LASSO-selected features, including duplicates, by training 50 LASSO models on random subsets of 80% of the training data. The size of this bag varied slightly depending on which subsets were chosen ($n = 2,350 \pm 80$). In the second stage, we used B to determine which feature components were most relevant to the classification task. This was done by examining the distribution of B over each feature component category. From these distributions, we selected a set C of the most frequently appearing 3 tumor regions, 3 MRI sequences, 4

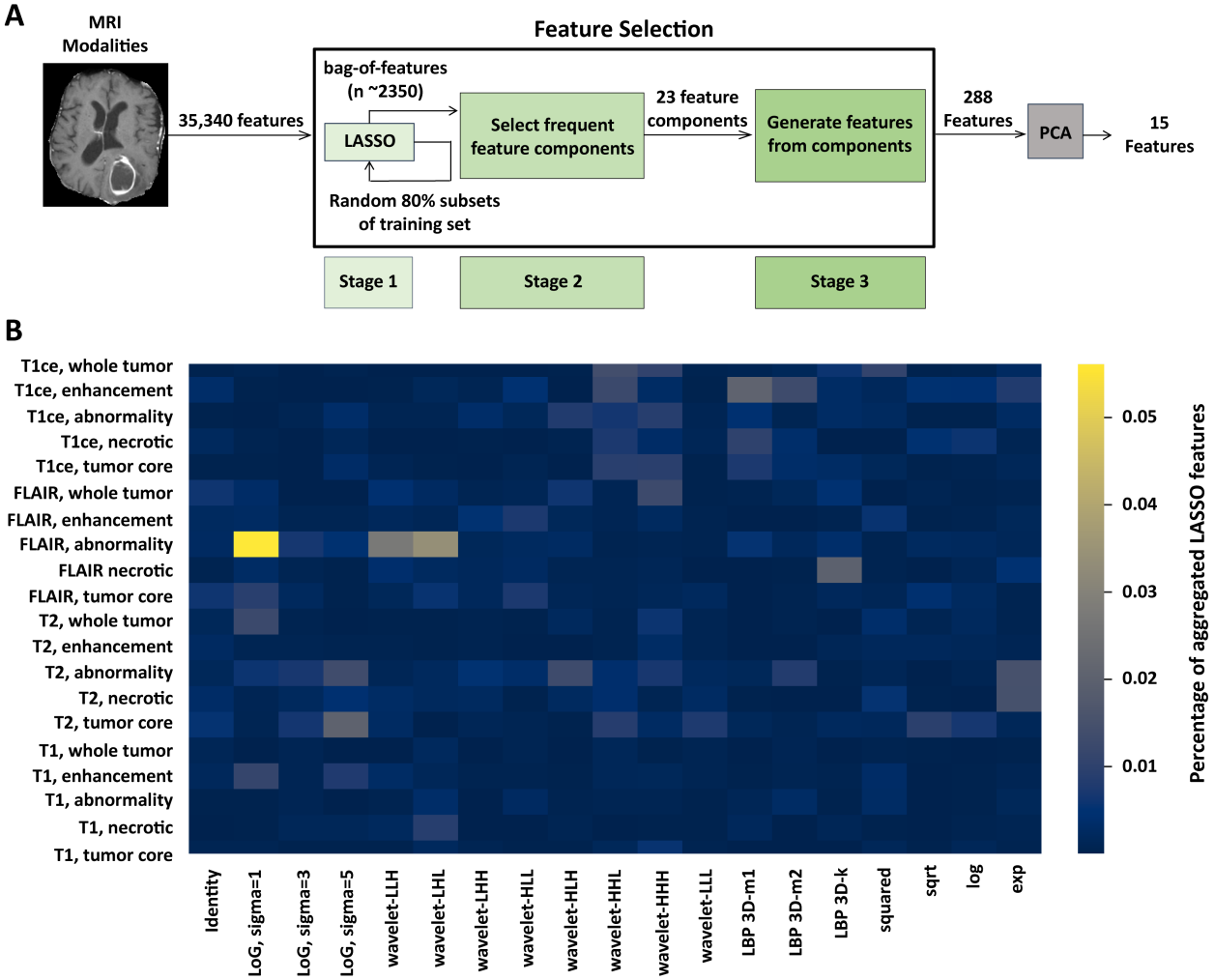


Figure 4.3: MRI feature selection schematic. (A) Feature selection pipeline starting with MRI extracted features and ending with 288 radiogenomic features. PCA further reduces the feature set to 15 principal components that are used to train machine learning models to predict patients' placement into IDH-wildtype molecular Group 1 versus Group 2. (B) Heatmap example of radiographic features selected in the first stage of our feature selection method. This shows the distribution of the features aggregated in the bag-of-features B over VOIs defined by a combination of MRI sequence, tumor region, and image transformation. Notably, the VOI defined by the FLAIR sequence, T2 abnormality, and LoG image transformation with kernel width 1 mm is the most common source of radiographic features in B.

image transformations, and 8 histogram/texture formulas, where thresholds were determined empirically. In the third stage, we generated the set of 288 features (Figure 4.3A) whose components were determined from the set C by selecting those radiographic features whose feature components were all contained in C. Finally, we used PCA to further reduce the dimensionality of our feature set to 15.

Statistical Methods: We trained and evaluated six classes of machine learning models including LASSO models, linear support vector machines (SVM), multi-layer perceptrons (MLP), XGBoost models, random forest models (RF), and ridge logistic regression (LR) classifiers on features selected by our proposed feature selection method (Python, Version 3.7, www.python.org). Model performance was measured by the area under the receiver operating characteristic curve (AUC). All reported numbers are the average AUC of 100 trials of 10-fold cross-validation run on randomly chosen partitions. Receiver operating characteristic (ROC) curves were also used to compare model performances. Feature selection, principal component analysis (PCA) dimensionality reduction, and model evaluation were conducted inside the cross-validation loop. To avoid overfitting, we used the default model hyperparameters from Python's scikit-learn and xgboost packages [110, 111]. All reported selected features were identified by conducting feature selection on the entire dataset after all parameters had been chosen.

4.4 Experiments and Results

Radiogenomic Model Prediction: To evaluate the efficacy of our feature selection method, we compared the results of machine learning models trained to bin patients into respective IDH-wildtype genomic groups with our feature selection method to the results of similar models trained with the following standard feature selection methods: LASSO feature selection, recursive feature elimination, and variance thresholding. For fair comparison across models, all feature selection methods were forced to select exactly 288 features and were evaluated with the application of 15-dimensional PCA reduction. Additionally, to determine

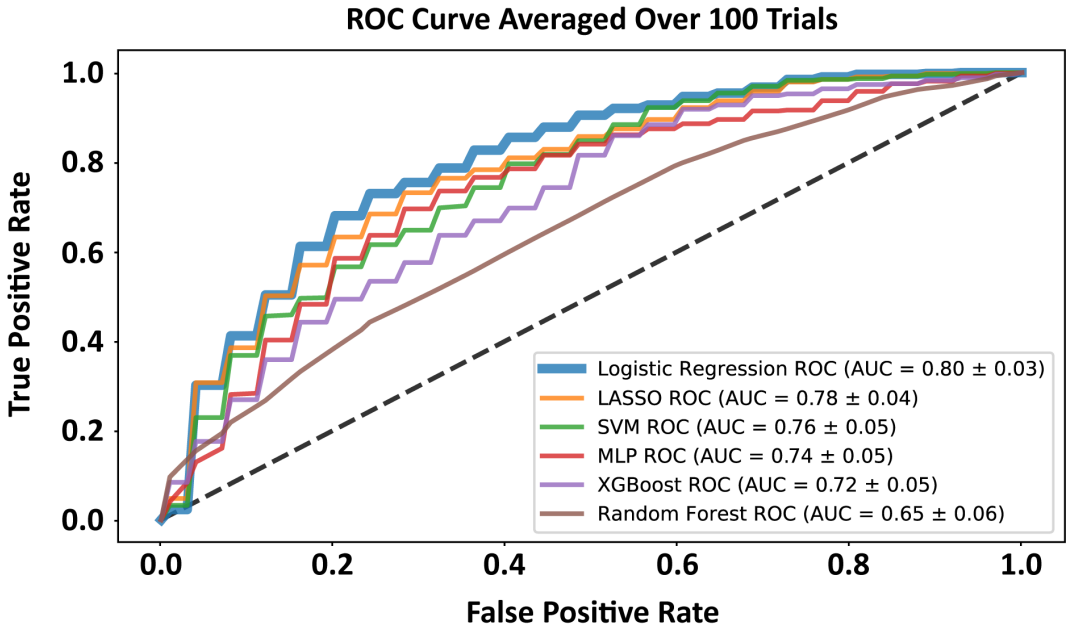


Figure 4.4: Comparison of Receiver Operating Characteristic (ROC) curves for various machine learning models that predict assignment to Group 1 versus Group 2 from MRI.

the utility of our feature selection method beyond reducing the dimension of model input, we compared our results to those of models trained on the 15-dimensional PCA embedding all 35,340 extracted radiographic features. Further, to better understand the extent to which molecular Group 1 and molecular Group 2 can be distinguished visually, we trained models on 60 tumor shape features and on 16 VASARI MRI features for comparison.

IDH-wildtype glioblastoma molecular groups 1 and 2 were classified with a cross-validated AUC score of 0.80 ($\pm$0.03) using a ridge logistic regression model trained with our PCA-reduced feature selection method (Figure 4.4). All models that were trained with our method significantly outperformed the same models trained with LASSO feature selection, recursive feature elimination, and variance thresholding (Table 4.1). Moreover, models trained with our feature selection method outperformed models trained on the 15-dimensional PCA embedding of all 35,340 features (Table 4.1). This comparison controls for model input dimen-

Features	**Lasso**	**SVM**	**MLP**	**XGBoost**	**RF**	**LR**
Ours	**0.78**	**0.76**	**0.74**	**0.72**	0.65	**0.80**
LASSO Feature Selection	0.59	0.58	0.63	0.56	0.54	0.61
Recursive Feature Elimination	0.58	0.58	0.61	0.55	0.54	0.59
Variance Thresholding	0.63	0.61	0.65	0.63	0.62	0.68
All Features (PCA = 15)	0.72	0.67	0.68	0.64	**0.71**	0.74
All Features	0.59	0.69	0.69	0.70	0.66	0.76
VASARI	0.42	0.46	0.41	0.59	0.48	0.43
Shape	0.53	0.38	0.53	0.40	0.59	0.51

Table 4.1: TCIA classification results (Mean AUC, 100 Trials)

sion and indicates that the improvement in performance provided by our feature selection method was due to selecting richer features rather than simply reducing the number of input features. Unlike models trained with our feature selection method, models trained with the standard feature selection methods we evaluated underperformed those trained on the PCA embedding of all extracted features. This supports the notion that our method is able to mitigate the risk of overfitting which is more common in other feature selection methods. The fact that no model trained on shape or VASARI features achieved an AUC score over 0.6 is a testament to the difficulty of distinguishing molecular Group 1 from molecular Group 2 with straightforward visual features. Examples of MRI volumes that were consistently correctly classified across cross-validation folds are shown in Figure 4.5 to showcase this difficulty and point out possible patterns.

We evaluated models and feature selection methods across a range of PCA dimension choices and chose the lowest round number (n=15) above which most models' performance

plateaued. Importantly, models trained with our feature selection method performed best across the range of choices of PCA dimension we evaluated (3-40), as well as when PCA was not applied. Interestingly, the results produced by models trained on feature selection methods without the application of PCA were mostly similar to those trained with PCA. Of note, our method was slightly more effective without the use of PCA (AUC=0.82 $\pm$0.03). Other changes induced by the application of PCA were decreased performances of random forest and XGBoost models trained on LASSO and recursive feature elimination, which were significantly higher when PCA was not applied, though their AUC scores did not exceed 0.70. These models also performed well when LASSO feature selection was permitted to select fewer than 288 features: random forest and XGBoost models achieved AUC scores of 0.76 ($\pm$0.06) and 0.78 ($\pm$0.07), respectively, when trained on small (n<10) sets of features selected by LASSO. Though these results were promising, performance was unstable and the features LASSO selected varied greatly between cross-validation folds, hindering interpretability. In general, models trained with LASSO feature selection, recursive feature elimination, and variance thresholding methods, even when they were not constrained to select 288 features, performed worse than models trained with our feature selection method.

Interpretation: The most interpretable selected histogram/texture formulas include histogram skewness and kurtosis and GLCM cluster shade, contrast, and informational measure of correlation. These features describe the symmetry and peakedness of an image-transformed VOI's intensity histogram and the uniformity, variation, and row-column correlation of the values in its grey-level co-occurrence matrix. Histogram symmetry and peakiness, as well as texture uniformity, may describe the diffusivity of gliomas which is linked to prognosis. The T1ce, FLAIR, and T2 MRI sequences were selected in the second stage of our feature selection method, as were the T2 abnormality, whole tumor, and tumor core tumor regions and two Laplacian of the Gaussian (LoG) and two Haar wavelet image transformations. The VOI defined by the FLAIR sequence, T2 abnormality region, and LoG image transformation with kernel width 1mm is likely the primary source of radiographic features that contribute most to the classification tasks because these feature components

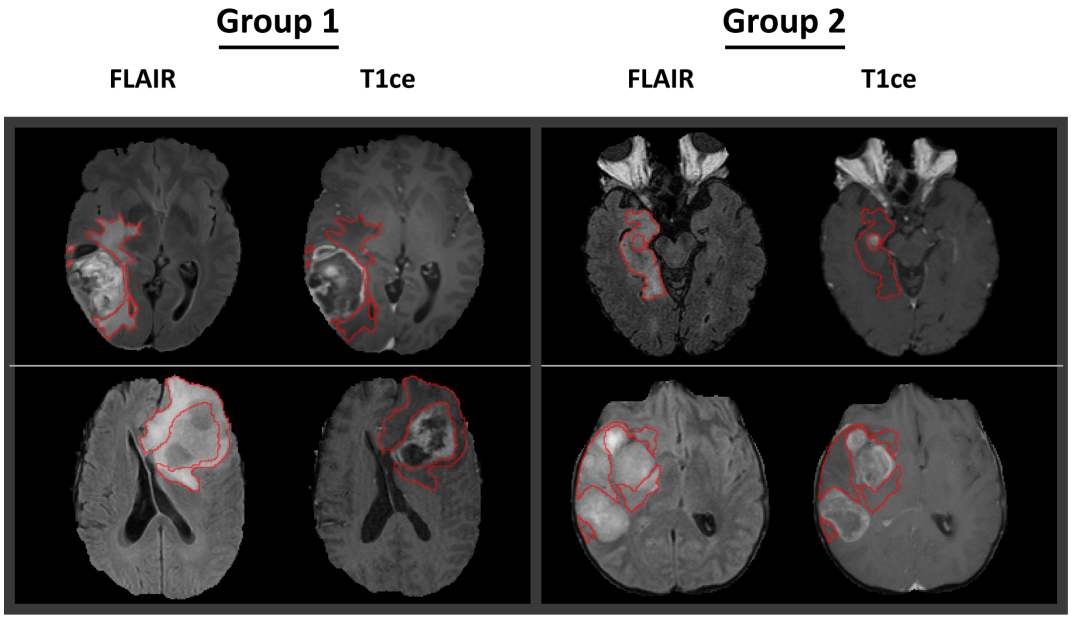


Figure 4.5: Examples of MRI volumes consistently correctly classified by ridge logistic regression. The T2 abnormality region boundary is outlined in red on the FLAIR and post-contrast T1ce images. Our results indicate that the T2 abnormality on the FLAIR sequence best discriminates between molecular Group 1 and molecular Group 2; however, this is not obvious to visual inspection. Sharply defined T2 abnormality region boundaries may be more apparent in Group 1, and more infiltrative T2 abnormality regions may be more common in Group 2.

were disproportionately represented in the bag B of aggregated features produced in the first stage of our feature selection method (Figure 4.3B). Moreover, because the LoG image transformation is an edge detector, we posit that the best discriminating MRI features describe the contours of hyperintensity within the transformed T2 abnormality region derived from the FLAIR sequence.

4.5 Discussion

We developed a novel LASSO-based feature selection method that improved the ability of machine learning classification algorithms to delineate prognostic molecular subgroups of IDH-wildtype glioblastomas originally defined by combined WES and SCNAs. In particular, a ridge logistic regression classifier trained on a 15-dimensional PCA embedding of 288 selected radiogenomic features predicted these two groups with cross-validated AUC=0.80 ($\pm$0.03). We additionally conjectured that the T2 abnormality tumor region in the FLAIR MRI sequence under a Laplacian of Gaussian (LoG) edge detector image transformation may be the primary source of the most discriminative signal in this task. We suspect this for the following two reasons. First, the T2 abnormality is the most prominent tumor region on the FLAIR sequence and both the T2 abnormality and FLAIR sequence were selected in the second stage of our feature selection method. Second, VOIs constructed from the FLAIR sequence and T2 abnormality were the most common source of features selected in the first stage of our method (Figure 4.3B). If true, fine-grain contours in the peritumoral edema and the infiltrating portion of glioblastomas on the FLAIR MRI sequence may be the basis for a characterization of the difference between molecular Group 1 and Group 2. However, this difference may be difficult for the human eye to parse. Were pronounced differences in FLAIR contours and infiltration visible, we would expect VASARI features to be better discriminators of Group 1 and Group 2 than we observed. Similarly, while examinations of correctly classified MRI samples may suggest that T2 abnormality region boundaries may be better defined in Group 1 and appear more infiltrative in Group 2, these observations are neither obvious nor conclusive (Figure 4.5). Additionally, the absence of the identity image transformation in the set of selected feature components may also attest to the difficulty of distinguishing between Group 1 and 2 by eye. We are more confident that tumor enhancement does not separate these groups. While our method selected the T1ce sequence, it did not select the enhancing tumor region, likely because the presence of tumor enhancement on the T1ce sequence of IDH-wildtype glioblastoma MRI scans is nearly

ubiquitous and thus may not differ significantly between molecular Group 1 and Group 2.

Our cross-validated results showed that models trained with our feature selection method outperformed baseline models trained without feature selection on the same 35,340-dimensional dataset and its 15-dimensional PCA embedding. On the other hand, models trained with LASSO feature selection, recursive feature elimination, and variance thresholding benchmarked well below these baseline results, a clear sign of overfitting. This indicates that our method is more robust to overfitting than three standard feature selection methods on a small set of high-dimensional data. The observation that linear models, such as logistic regression and LASSO, outperformed more complex models, such as random forests and support vector machines, shows the benefit of using simple models in such settings to avoid overfitting in the post-feature selection modeling phase.

This study's contribution to radiogenomics is a multi-stage feature selection method that boosted the performance of a set of machine learning classifiers in a situation where the number of features significantly outnumbered the number of training samples. Such situations are likely to remain commonplace in medical imaging as the accrual of pre-operative MRI data is constrained by the limited cases of glioblastoma [5]. Further, with the inception of open-source radiogenomic feature extraction software, such as *pyradiomics* [100] and the Cancer Imaging Phenomics Toolkit (CaPTk) [112], the MRI sequence, tumor region, and imaging transformation-based radiogenomic feature structure that our method leverages is increasingly becoming the standard, thereby ensuring the relevance of our method.

Moving forward, additional cohorts will be necessary to rigorously validate our method. However, currently, there is a lack of available datasets for patients with de novo IDH-wildtype glioblastomas that contain WES, SCNA data, and multiple MRI sequences. Nonetheless, if our method's robustness is further supported outside of our current cross-validation study, such an MRI-based classifier capable of distinguishing poor surviving Group 1 IDH-wildtype glioblastoma patients from longer surviving Group 2 IDH-wildtype glioblastoma patients could have an immediate influence on patient care. Shorter-term Group 1 patients, who are also less likely to have a surgical intervention for tumor recurrence [15], could be

recommended for upfront clinical trials. Within trials, if Group 1 and Group 2 patients are not balanced in phase II and phase III arms, incorrect and costly conclusions may be drawn. Inferring patient risk stratification from available baseline data such as MRI is a critical way forward in neuro-oncology.

Chapter 5

LEVERAGING UNLABELED DATA FOR GLIOMA MOLECULAR SUBTYPE AND SURVIVAL PREDICTION

Abstract

In this chapter, we address two long-standing radiogenomic challenges in glioma subtype and survival prediction: (1) how to leverage large amounts of unlabeled magnetic resonance imaging (MRI) data and (2) how to unite MRI data and genomic data. We propose a novel application of multi-task learning (MTL) that leverages unlabeled MRI data by jointly learning an auxiliary tumor segmentation task with glioma subtype prediction and that can learn from patients with and without genomic data. We analyze multi-parametric MRI data from 542 patients in the combined training, validation, and testing sets of the 2018 Multimodal Brain Tumor Segmentation Challenge and somatic copy number alteration (SCNA) data from 1090 patients in The Cancer Genome Atlas' (TCGA) lower-grade glioma and glioblastoma projects. Our MTL model significantly outperforms comparable classification models trained only on labeled MRI data for both IDH1/2 mutation and 1p/19q co-deletion subtype prediction tasks. We also show that embeddings produced by our MTL models improve survival predictions beyond MRI or SCNA on their own. Our code is available at `https://github.com/nknuecht/glioma_mtl`.

5.1 Motivation

IDH mutation and 1p/19q co-deletion status are the most important molecular markers for adult diffuse glioma. Knowledge of the status of these alterations prior to surgery may influence the aggressiveness of the surgeon's approach. In the extreme case where neither a resection nor biopsy can be conducted, medical imaging data are the only data from which

the presence or absence of an IDH mutation or 1p/19q co-deletion can be inferred. Machine learning models have been used to predict IDH mutations and 1p/19q co-deletions from magnetic resonance imaging (MRI) in the past, but their results are difficult to compare and have not taken advantage of adult diffuse glioma MRI data without IDH mutation or 1p/19q co-deletion information. Further, these models do not incorporate genomic data in the cases where it is available. Leveraging unlabeled imaging data and available genomic information can recruit data otherwise unused and may improve prediction results.

5.2 Related Literature

Radiogenomics has advanced rapidly in the last decade, but most methods remain challenging to compare [23]. Many studies are conducted on private datasets, and code for radiogenomic pipelines is not always released [113]. In contrast, the Multimodal Brain Tumor Segmentation Challenge (BraTS) has spearheaded open science in glioma tumor segmentation by curating large, multi-parametric glioma MRI datasets [103, 102, 34]. The medical imaging and computer vision communities have used the BraTS datasets to train state-of-the-art segmentation models and build valuable open-source software tools [34, 114, 112]. Similarly, public genomic data in the TCGA have been used for groundbreaking, reproducible research, especially in glioma [38, 40, 39, 42, 45]. Of particular interest are somatic copy number alteration (SCNA) data, which describe large, contiguous deletions or duplications of DNA and have been linked to glioma molecular subtype and survival [46, 15].

Importantly, the set of patients in the 2018 BraTS dataset partially overlaps with the set of glioma patients in the TCGA [115, 116]. Labels for IDH mutations and 1p/19q co-deletions as well as overall survival (OS) are available for most patients in the intersection of these datasets, as is SCNA data. However, the majority of BraTS patients do not have accompanying data in the TCGA. Thus, training models only on BraTS MRI data with genomic or survival labels leaves the bulk of these data unused, and training models only on imaging data ignores rich genomic data when it is available.

Multi-task learning (MTL) is a machine learning strategy that allows models to learn mul-

tiple tasks at once and is ideal for incorporating different but related data sources [117, 118]. Compared to single-task learning, MTL allows models to learn richer data representations from more diverse data and labels by receiving feedback from complementary tasks, such as tumor segmentation and IDH mutation and 1p/19q co-deletion prediction [119, 120].

5.3 Problem Statement

In this paper, we address two long-standing radiogenomic challenges in adult diffuse glioma subtype and survival prediction: (1) how to leverage large amounts of unlabeled magnetic resonance imaging (MRI) data and (2) how to unite MRI data and genomic data. We propose a novel application of multi-task learning (MTL) that leverages unlabeled MRI data by jointly learning an auxiliary tumor segmentation task with IDH mutation and 1p/19q co-deletion status prediction and that can learn from patients with and without somatic copy number alteration (SCNA) data.

5.4 Datasets

Multi-parametric MRI data were downloaded for 542 patients from the 2018 BraTS training, validation, and testing datasets, as described in Section 2.1.2. These MRI data are 4-channel volumes composed of pre- (T1) and post-contrast (T1ce) T1-weighted sequences and T2-weighted (T2) and T2 Fluid-Attenuated Inversion Recovery (FLAIR) sequences (Figure 5.1A). For 285 of the 542 BraTS patients, 4-class segmentation masks that denote three different tumor compartments and a background class (Figure 5.1B) are given. Gene-level SCNA data were downloaded for 1090 patients in the TCGA lower-grade (WHO grade 2/3) glioma and glioblastoma (WHO grade 4) projects from the University of California Santa Cruz cancer browser. All 235 patients in the intersection of the BraTS MRI dataset and the TCGA glioma SCNA dataset have overall survival (OS) and 1p/19q co-deletion labels, while only 171 have IDH mutation labels. We use the 75 patients in the BraTS 2018 validation and testing sets with SCNA data as a validation set for our models, of which 59 have IDH labels. Of the 307 samples in the BraTS dataset with neither survival nor subtype labels,

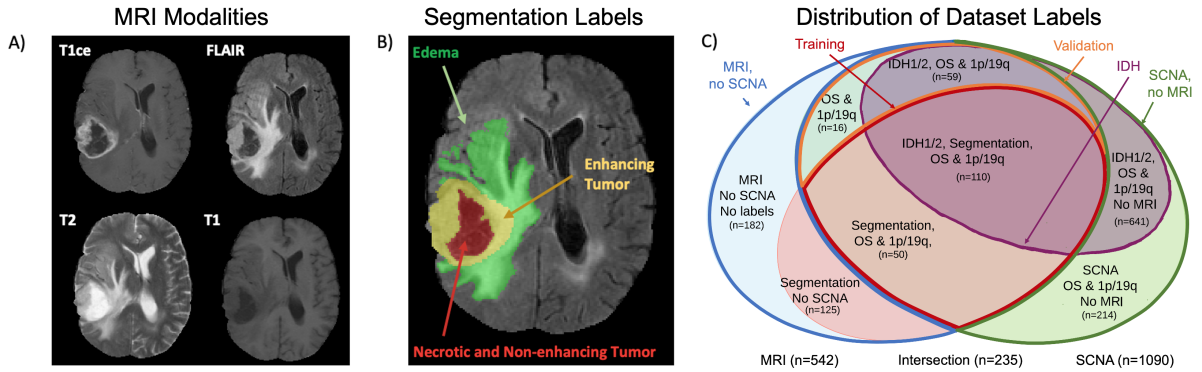


Figure 5.1: **A)** Slices of four MRI modalities of a WHO grade IV IDH1/2 wildtype tumor in the BraTS dataset. Each MRI sample has dimension $4 \times 240 \times 240 \times 155$, where the 4 channels represent the T1ce, FLAIR, T2, and T1 MRI modalities. The hallmark enhancing ring of aggressive tumors is clearly visible on the T1ce modality. **B)** A ground truth BraTS 4-class segmentation mask overlaid on the FLAIR modality. The outer class consists of peritumoral edema; the inner classes consist of necrotic tissue, non-enhancing tumor, and enhancing tumor. **C)** Distribution of dataset labels. The labeled training set is outlined with a red boundary; the validation set is outlined in an orange boundary; the unlabeled MRI data leveraged by our MTL model is outlined with a blue boundary; the set of samples with SCNA data but without MRI data is outlined with a green boundary. Notably, while all samples in the labeled training set and validation set have 1p/19q co-deletion and survival labels, only 69% of samples in the labeled training set have IDH labels, and only 78% of the samples in the validation set have IDH labels. Additionally, while we refer to the MRI samples without subtype markers or survival labels as unlabeled, 41% of these samples do have 4-class segmentation labels.

125 have ground truth segmentation labels. We assign weak segmentation labels to the other 182 samples using a tumor segmentation model we trained on the BraTS training data [57]. The distribution of labels in our dataset is shown in Figure 5.1C.

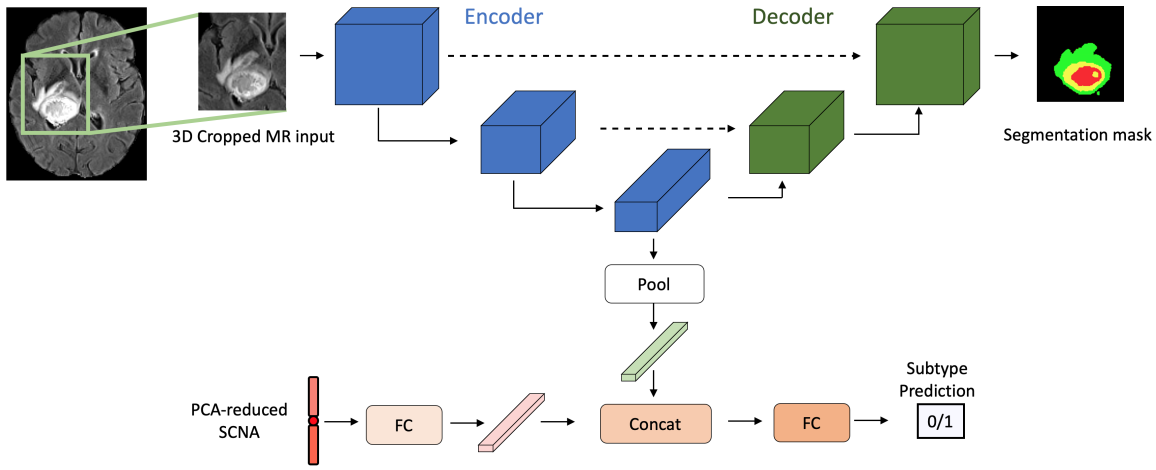


Figure 5.2: Our MTL model uses the architecture of 3D-ESPNet with a classification branch connected to the output of the encoder. PCA-reduced genomic SCNA data is passed through fully connected layers, concatenated with the average-pooled encoder output, and fed into a classifier to predict subtype class. The network decoder also produces a segmentation mask. We take the weighted cross-entropy loss of the subtype classification and tumor segmentation tasks. Our model accepts full brain multi-model MRI volumes as well as cropped tumor volumes (shown here). We train models on 4-channel MRI data input and 1-channel, single-sequence input.

5.5 Methods

Model: Our MTL model is illustrated in Figure 5.2. The backbone of our network is an open-source ESPNet-based [90] U-Net style [121] segmentation network we trained on the BraTS 2018 training dataset [57]. To predict molecular subtype markers, we add a branch to the bottom of the network by average pooling the output of the network's encoder. To allow SCNA data to contribute to prediction, we pass 50-dimensional SCNA PCA embeddings through a small fully connected network and concatenate its output with the output of the average pooling step. We then pass this concatenated vector through a fully connected layer to obtain a binary IDH mutation or a binary 1p/19q co-deletion prediction. Notably, these

PCA embeddings are generated from all 1,015 SCNA samples not included in the validation set (Figure 5.1, outlined in red and green). Thus, we are able to leverage TCGA data that does not overlap with the BraTS dataset.

Loss Function: For subtype classification, we take the weighted binary cross-entropy loss $\mathcal{L}_C$ for training samples that have subtype labels (Figure 5.1, outlined in red) in addition to a variable segmentation loss. For training samples with 4-class ground truth segmentation labels (Figure 5.1, red oval partially outlined in red), we take the weighted 4-class cross-entropy segmentation loss $\mathcal{L}_{S_{gt}}$ between the MTL model's decoder's output and the supplied segmentation mask. For unlabeled MRI training samples for which 4-class segmentation labels are not available (Figure 5.1, outer blue crescent), we take the weighted 2-class cross-entropy segmentation loss $\mathcal{L}_{S_{weak}}$ between the binarized output of the MTL model's decoder and the binarized weak segmentation mask supplied by the pre-trained segmentation network. We binarize our weak segmentation masks, because whole tumor segmentation is known to be more reliable than within-tumor region segmentation for BraTS pre-trained models [57, 34]. Finally, we define our MTL loss as

$$\mathcal{L} = \mathcal{L}_C + \lambda \mathcal{L}_{S_{gt}} + (1 - \lambda)\mathcal{L}_{S_{weak}} \tag{5.1}$$

where $\mathcal{L}_C = 0$ for samples without subtype labels, and λ controls the feedback from weak and ground truth segmentation labels. For samples with ground truth segmentation labels, we set $\lambda = 1$. Otherwise, we set $\lambda = 0$.

We also train our MTL model on 1-channel, single-sequence MRI input to assess the predictive power of each MRI sequence. When we train on T1-weighted sequences (T1ce, T1, T1ce-T1), we do not consider the edema segmentation label (Figure 5.1B) when evaluating $\mathcal{L}_{S_{gt}}$ because it is difficult to distinguish edema on these sequences. When we train on the T2 and FLAIR sequences, we binarized the tumor segmentation labels, because these sequences characterize the tumor boundary better than any of the interior compartments. We do not modify $\mathcal{L}_{S_{weak}}$.

5.6 Experiments and Results

Experiments: We first establish that utilizing unlabeled MRI data boosts glioma subtype classification performance. To do this, we compare our MTL model trained with unlabeled MRI data to 3D CNNs trained only on labeled MRI data. For a fair comparison, we match the architecture of the CNNs with that of our MTL model's encoder and use the same hyperparameters. Second, we show that the performance of our MTL model improves when we allow for SCNA input. Third, we show that MRI and SCNA data predict survival better than either on their own. To do this, we train linear Cox proportional hazards (CPH) models on embeddings produced by MTL models trained on (1) MRI data alone and (2) MRI and PCA-reduced SCNA data and compare their results to those of a linear CPH model trained directly on PCA-reduced SCNA data. We perform these experiments with the following MRI input formats:

- *All sequences (whole brain):* 4-channel volume consisting of all four MRI sequences.
- *All sequences (cropped):* 4-channel volume consisting of all four MRI sequences cropped to a tumor bounding box found using either ground truth or weak segmentation labels.
- *Single sequences (cropped):* Cropped 1-channel volumes consisting of a single MRI sequence.
- *T1ce-T1 sequence (cropped):* Cropped 1-channel cropped volume constructed by subtracting the T1 volume from the T1ce volume. This volume accentuates the hallmark enhancing tumor region found on the T1ce sequence (Figure 5.1A,B).

Evaluation Metrics and Training Details: For the IDH mutation and 1p/19q co-deletion prediction tasks, we report the average maximum area under curve (AUC) score for each model trained for 50 epochs over 10 trials. For survival prediction, we report the C-index, which measures the extent to which a model can properly order survival time. These are

Input Sequences	IDH Mutation (AUC)		1p/19q Co-deletion (AUC)		Overall Survival (C-index)	
	CNN	MTL (MRI)	CNN	MTL (MRI)	MTL (MRI)	MTL (MRI+SCNA)
All (Whole Brain)	0.669	**0.846**	0.605	**0.813**	0.587	**0.723**
All (Cropped)	0.872	**0.894**	0.744	**0.871**	0.697	**0.732**
T1ce (Cropped)	**0.893**	0.884	0.772	**0.819**	0.719	**0.735**
FLAIR (Cropped)	**0.778**	0.690	0.755	**0.818**	0.565	**0.731**
T1 (Cropped)	0.731	**0.738**	0.727	**0.757**	0.645	**0.728**
T2 (Cropped)	**0.778**	0.732	0.740	**0.755**	0.690	**0.718**
T1ce-T1 (Cropped)	**0.895**	0.861	**0.764**	0.742	0.707	**0.723**

Table 5.1: Results comparing MTL models across prediction tasks and MRI input format.

standard metrics in glioma MRI-based classification and survival prediction [122, 30, 123, 28, 29, 124].

We split our MRI dataset into 467 labeled and unlabeled MRI training samples (Figure 5.1, large oval outlined in blue and green) and 75 MRI validation samples (Figure 5.1, outlined in orange) defined by the intersection of the BraTS validation and testing datasets with the TCGA glioma SCNA dataset. We use the remaining 855 samples with SCNA data but no MRI data (Figure 5.1, crescent outlined in green) to improve our SCNA PCA embeddings.

IDH Mutation and 1p/19q Co-deletion Results: Table 5.1 shows IDH mutation and 1p/19q co-deletion prediction results for models trained on the set of MRI inputs detailed previously. The most dramatic boost MTL models give is on whole brain, all sequence input, where they raise classification AUC by 0.18 for IDH mutation prediction and 0.21 for 1p/19q co-deletion prediction. Our MTL models almost ubiquitously outperform CNNs on the 1p/19q co-deletion task, but the results on the IDH task are less clear. We suspect that our MTL models do not significantly improve IDH prediction because both our best CNNs and MTL models are able to learn that tumor enhancement is strongly associated with IDH wildtype tumors: in the labeled training set, 85.5% of IDH wildtype tumors

Input	1p/19q Co-deleted		IDH Mutant, 1p/19q Intact		IDH Wildtype	
	MRI	**MRI + SCNA**	**MRI**	**MRI + SCNA**	**MRI**	**MRI + SCNA**
All (Whole Brain)	0.714	**1.000**	0.606	**0.727**	0.487	**0.521**
All (Crop)	0.607	**0.964**	**0.742**	0.712	0.540	**0.548**
T1ce (Crop)	**0.821**	0.786	0.576	**0.742**	**0.644**	0.571
FLAIR (Crop)	0.607	**0.964**	0.636	**0.712**	0.527	**0.540**
T1 (Crop)	0.643	**0.893**	0.606	**0.636**	0.535	**0.544**
T2 (Crop)	**0.679**	0.500	**0.803**	0.697	0.562	**0.563**
T1ce-T1 (Crop)	0.821	**0.857**	**0.803**	0.682	0.523	**0.552**
SCNA, PCA=5	0.929		0.667		0.512	

Table 5.2: MTL survival performance broken up over WHO 2016 glioma subtypes (C-index).

show strong enhancement and 83.6% of IDH mutant tumors show mild or no enhancement. A visualization of integrated gradients [125] and examples of incorrect IDH predictions in Figure 5.3 corroborate this explanation. Last, all MTL models trained with MRI and SCNA data were able to classify IDH mutation and 1p/19q co-deletion status with AUC > 0.98. We leave these results out of Table 5.1 because they are too similar to compare.

Survival Results: The last columns of Table 5.1 show that embeddings produced by MTL models trained on SCNA and MRI data better predict survival than embeddings produced by MTL models trained on MRI data alone. In Table 5.2, we break up our survival prediction results by WHO 2016 molecular subtype[1] and observe that these models perform exceptionally well on 1p/19q co-deleted tumors and offer some improvement for IDH wild-type gliomas. Our best MTL-embedding-based survival models also compare favorably to linear CPH models trained on PCA-reduced SCNA data alone, suggesting adding MRI data to SCNA data improves survival prediction.

[1]All gliomas with 1p/19q co-deletions have IDH mutations.

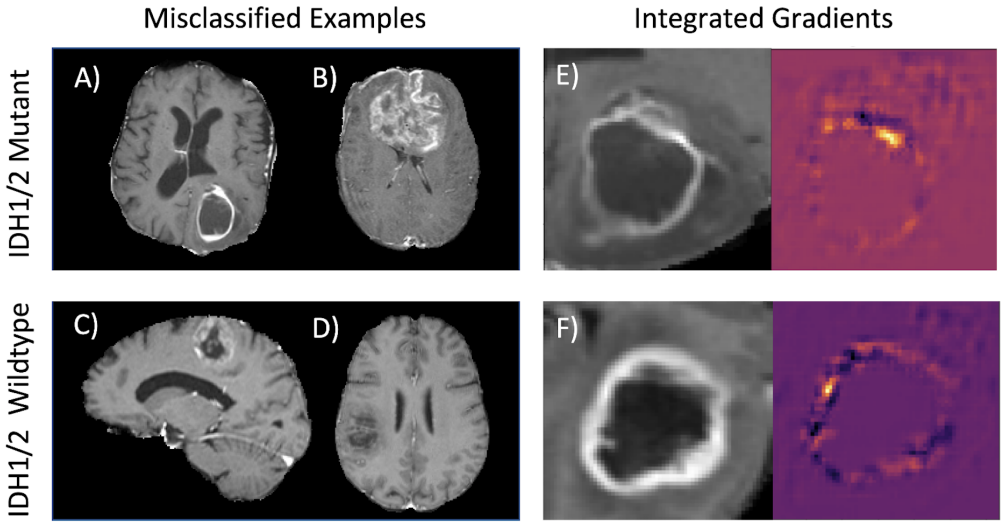


Figure 5.3: Our T1ce-based IDH MTL network appears to associate tumor ring enhancement with IDH wildtype tumors. Sample **A,B, E)** are IDH mutant glioma with ring enhancement misclassified as IDH wildtype tumors. Samples **C,D)** are IDH wildtype tumors with mild and no enhancement misclassified as IDH mutants. Sample **F)** is a correctly classified IDH wildtype. Integrated gradients in the images to the right of **E)** and **F)** show that this model puts emphasis on tumor ring enhancement.

5.7 Discussion

The primary contribution of this study is a novel, reproducible application of MTL that (1) jointly learns tumor segmentation with IDH or 1p/19q co-deletion subtype so that samples without subtype labels can contribute to model learning and (2) unites glioma genomic and imaging data by allowing SCNA data to serve as model input alongside MRI data. Our MTL model that leverages unlabeled, 4-channel MRI data accurately predicts IDH mutations (AUC = 0.89) and 1p/19q co-deletions (AUC = 0.87) and outperforms comparable 3D CNNs trained on labeled MRI imaging data alone. Training with SCNA data dramatically boosts subtype performance (AUC $>$ 0.98) and raises the survival C-index score of survival models trained on embeddings generated by IDH prediction models from 0.719 to 0.735. To interpret

our results, we trained and evaluated models on individual MRI sequences, and visualized model gradients, and interrogated the relationship between our models' predictions. To the best of our knowledge, this is the first multi-task learning strategy that unites the BraTS MRI and TCGA glioma SCNA datasets.

The observation that adding SCNA input to our MTL models drastically improves both subtype tasks (AUC > 0.98) across all MRI inputs is notable. This is likely explained by the facts that 1p/19q co-deletions are deductible from SCNA data and IDH wildtype tumors commonly carry distinctive SCNAs such as whole gain of chromosome 7 and whole loss of chromosome 10 [12]. On the other hand, links between SCNA data and patient survival are not as straightforward, and thus it is impressive that the inclusion of imaging data in our MTL-embedding-based survival models improves results set by survival models trained on SCNA data alone.

We emphasize the clinical applications of this study. For patients who are unable to undergo brain surgery, or otherwise cannot obtain their IDH mutation and 1p/19q co-deletion status, accurate MRI-based predictions of subtype place patients on different survival trajectories. For patients whose IDH mutation status and 1p/19q co-deletion status are known, our survival models offer the potential for sub-stratifying survival within glioma subtype. Improvements to survival stratification may lead to better treatment management, especially for predicted short-term survivors for whom early clinical trial enrollment may be recommended.

Chapter 6

MACHINE LEARNING MODELING OF GENOME-WIDE COPY NUMBER ALTERATION SIGNATURES RELIABLY PREDICTS IDH MUTATIONAL STATUS IN ADULT DIFFUSE GLIOMA

Abstract

Knowledge of 1p/19q-codeletion and *IDH1/2* mutational status is necessary to interpret any investigational study of diffuse gliomas in the modern era. While DNA sequencing is the gold standard for determining IDH mutational status, genome-wide methylation arrays and gene expression profiling have been used for surrogate mutational determination. Previous studies by our group suggest that 1p/19q-codeletion and IDH mutational status can be predicted by genome-wide somatic copy number alteration (SCNA) data alone; however, a rigorous model to accomplish this task has yet to be established. In this study, we used SCNA data from 786 adult diffuse gliomas in The Cancer Genome Atlas (TCGA) to develop a two-stage classification system that identifies 1p/19q-codeleted oligodendrogliomas and predicts the IDH mutational status of astrocytic tumors using a machine-learning model. Cross-validated results on TCGA SCNA data showed near-perfect classification results. Furthermore, our astrocytic IDH mutation model validated well on four additional datasets (AUC=0.97, AUC=0.99, AUC=0.95, AUC=0.96) as did our 1p/19q-codeleted oligodendroglioma screen on the two datasets that contained oligodendrogliomas (MCC=0.97, MCC=0.97). We then retrained our system using data from these validation sets and applied our system to a cohort of REMBRANDT study subjects for whom SCNA data, but not IDH mutational status, is available. Overall, using genome-wide SCNAs, we successfully developed a system to robustly predict 1p/19q-codeletion and IDH mutational

status in diffuse gliomas. This system can assign molecular subtype labels to tumor samples of retrospective diffuse glioma cohorts that lack 1p/19q-codeletion and IDH mutational status, such as the REMBRANDT study, recasting these datasets as validation cohorts for diffuse glioma research.

6.1 Introduction

Diffuse gliomas comprise the most common adult malignant tumors of the central nervous system [126]. These adult diffuse gliomas consist of three major biologically and clinically distinct molecular subtypes, which are defined by the mutational status of *isocitrate dehydrogenase 1 and 2* (IDH) and the presence or absence of co-deletion of whole chromosome arms 1p and 19q, which further stratifies IDH-mutant diffuse glioma [5, 6]. These genetic alterations are strong predictors of survival and contain more information than historical histologically-based classification and grading systems [42, 127, 128, 129, 130]. In routine surgical neuropathology, it is common practice to classify diffuse gliomas in terms of IDH mutational and 1p/19q-codeletion status. Likewise, in contemporary research studies of adult diffuse gliomas, there is limited utility, if any, in gaining insights into the biology of gliomas if the samples are not well annotated for this molecular information. Furthermore, determining the robustness and reliability of any findings in human gliomas requires testing and validation across multiple cohorts. Consequently, older retrospective cohorts of diffuse gliomas lacking IDH mutational and 1p/19q-codeletion status have limited utility for validating contemporary adult diffuse glioma study results.

As a testing platform, DNA sequencing is the gold standard method to detect the spectrum of clinically relevant canonical and non-canonical IDH mutations [131, 132, 133, 134, 135]. More recently, methods have been developed to infer IDH mutation and 1p/19q-codeletion status from methylation array [47, 39], gene expression [45], and magnetic resonance imaging data [27, 136, 29, 26], which can provide surrogate molecular subtype labels for validating adult diffuse glioma study results. We have previously observed that there is a strong association between adult diffuse glioma molecular subtype and patient

somatic copy number alteration (SCNA) profiles, indicating that SCNA data alone may reflect global genomic structures that are associated with, and predictive of, IDH mutational status [95, 15, 46, 16, 137]. Furthermore, SCNA data has the advantage of directly encoding the extent of 1p and 19q loss, although an empirical threshold necessary to definitively call 1p/19q-codeletions has yet to be established.

In this study, we sought to develop and evaluate a robust system that predicts adult diffuse glioma IDH mutational status and 1p/19q-codeletion status from SCNA data alone. Special care is given to establish appropriate thresholds for calling 1p/19q-codeletions as well as simultaneous gain of whole chromosome 7 and loss of whole chromosome 10 (+7/-10), the latter of which is necessary for molecular grading of IDH-wildtype diffuse astrocytic gliomas [12]. We validate our system on a TCGA holdout set of histological World Health Organization (WHO) grade 4 tumors and three additional independent diffuse glioma datasets, including a dataset published by Glioma Longitudinal AnalySiS Consortium (GLASS) [74, 47, 75, 73]. Finally, we report the system's predictions on the retrospective REMBRANDT study, where genome-wide SCNA data is available, but IDH sequencing is not [76]. Overall, our study makes older adult diffuse glioma datasets with SCNA data but without molecular diagnoses better suited for validating contemporary findings. Additionally, this study proposes evidence-based thresholds for 1p/19q-codeletions and +7/-10.

6.2 Methods

TCGA Glioma Dataset

Somatic mutation calls for TCGA gliomas (N=812) computed by the Multi-Center Mutation Calling in Multiple Cancers (MC3) project [83] were obtained as described in Section 2.5.2. We obtained three versions (UCSC hg19, GDC hg19, GDC hg38) of the TCGA SCNA dataset as described in Section 2.5.2.

TCGA Glioma Ultra-Low-Pass Whole Genome Sequencing Dataset

TCGA data was processed as ultra-low-pass whole genome sequencing (ULP-WGS, 0.1x) to compare Affymetrix SNP 6.0 array-derived SCNA data to lower coverage data. Bam files for

every patient in the TCGA-LGG and TCGA-GBM projects with available bam files were downloaded from the GDC. These bam files were realigned to hg19 using SAMtools [138], BEDTools [139], Bowtie 2 [140], the Picard Toolkit (version 2.7.1), and the Genome Analysis Toolkit (GATK, version 3.7) [141]. The R package HMMcopy was used to create wig files, and the R package ichorCNA [142] was used to create a custom panel of normals from bam files of normal blood or tissue downloaded from GDC for patients in the TCGA-LGG and TCGA-GBM projects. ichorCNA was then used to create copy number segmentation files using wig files created by HMMcopy. GISTIC was used to compute gene-level SCNA calls from these segmentation files.

TCGA Molecular Subtype Training Labels

Our baseline IDH mutation and 1p/19q-codeletion training labels are derived from IDH mutation calls reported by the MC3 project [83] and 1p/19q-codeletion annotations published by Ceccarelli et al. [45]. We compared the MC3 IDH mutation labels to the IDH mutation labels published by Ceccarelli et al. and the 1p/19q-codeletion status published by Ceccarelli et al. to the gene-level TCGA SCNA training data. Two patients (TCGA-06-0151, TCGA-HT-A618) called IDH-wildtype in the MC3 project are labeled as IDH-mutant astrocytomas by Ceccarelli et al. One patient (TCGA-06-0151) is histologically WHO grade 4, harbors +7/-10, and lacks mutations in TP53 and ATRX. Because these are all characteristics of IDH-wildtype glioblastoma, we maintain the MC3 IDH-wildtype designation. The other patient (TCGA-HT-A618), however, is histological WHO grade 3, has intact chromosomes 7 and 10, and harbors TP53 and ATRX mutations; thus, we replace the MC3 project IDH-wildtype label with the Ceccarelli et al. IDH-mutant astrocytoma designation. Conversely, one patient (TCGA-P5-A72U) that carries an IDH mutation in the MC3 project data is labeled IDH-wildtype by Ceccarelli et al. This patient harbors +7/-10, and lacks mutations in TP53 and ATRX, and we, therefore, consider this an IDH-wildtype glioblastoma. While SCNA data from three TCGA SCNA pipelines indicate all 171 Ceccarelli et

al. oligodendroglial designated tumors carry 1p/19q-codeletions, five Ceccarelli et al. designated IDH-mutant astrocytomas carry significant loss of 1p and 19q without loss of either 1q or 19p on at least one, but not all, of the three TCGA SCNA versions we used (UCSC hg19, GDC hg19, GDC hg38) (Supplemental Figure 1). Furthermore, all five tumors are TP53-wildtype, ATRX-wildtype, and are classified histopathologically as oligodendroglioma or mixed oligoastrocytoma, indicating that they may be oligodendrogliomas. However, to avoid uncertainty, we excluded these five tumors from this study.

TCGA Training Set

Our final TCGA training set consists of 786 adult diffuse gliomas, 171 of which are IDH-mutant and 1p/19q-codeleted oligodendrogliomas, 257 of which are IDH-mutant astrocytomas, and 358 of which are IDH-wildtype glioblastomas. Apart from the five Ceccarelli et al. designated IDH-mutant tumors we justified excluding above, we also excluded 21 histological lower-grade IDH-wildtype tumors, because we could not confirm that they harbored simultaneous gain of whole chromosome 7 and loss of whole chromosome 10 (+7/-10), EGFR amplification, or TERT promoter (TERTp) mutation required by the fifth edition of the WHO criteria for classification as adult diffuse gliomas [12, 6]. The exclusion of these tumors is further justified in the Results section. The remaining 70 histological lower-grade IDH-wildtype tumors showed molecular markers consistent with the updated definition of IDH-wildtype glioblastoma as described in the fifth edition of the WHO criteria for classification as adult diffuse gliomas.

TCGA Validation Set

In addition to the 786 adult diffuse gliomas in our TCGA training set and the 26 tumors we excluded from our study, 167 other TCGA tumors have IDH mutational and 1p/19q-codeletion status annotations published by Ceccarelli et al., although they do not have MC3 mutational data to confirm these annotations. This cohort of patients consists of 2 oligodendrogliomas, 17 histological grade 4 IDH-mutant astrocytomas, 146 histological grade 4

IDH-wildtype glioblastomas, and 2 histological lower-grade IDH-wildtype tumors that do not qualify as adult diffuse IDH-wildtype glioma under the criteria in the fifth edition of the WHO classification of CNS tumors [6, 19]. We restrict this cohort to the 163 grade 4 astrocytic tumors and use these patients as a validation dataset.

Independent Validation Glioma Datasets

The Glioma Longitudinal AnalySiS (GLASS) Data

We used the Synapse API to download copy number segmentation files (variants-gatk-seg) for 201 primary diffuse astrocytic gliomas in the Glioma Longitudinal AnalySiS (GLASS) dataset (Data Release version 2019-03-28) [74, 75]. We processed these data with GISTIC with the parameters described below. This patient cohort consisted of 143 IDH-wildtype glioblastomas and 58 IDH-mutant astrocytomas determined from available IDH mutation annotations. We did not use data from GLASS oligodendroglioma patients because there were too few. Patients' clinical variables including overall survival (N=184) and age (N=186) were also downloaded.

Jonsson et al. Data

We obtained SCNA data for 432 primary diffuse glioma patients originally studied in Jonsson et al. as described in Section 2.2.2 [73]. This patient cohort consisted of 319 IDH-wildtype gliomas, 63 IDH-mutant astrocytomas, and 50 oligodendrogliomas. Published clinical variables, including overall survival (N=432) and age (N=432) for all patients, were also downloaded.

Capper et al. Data

We obtained SCNA data analyzed in Capper et al. as described in Section [47]. This patient cohort consisted of 298 IDH-wildtype glioblastomas, 110 IDH-mutant astrocytomas, and 81 IDH-mutant and 1p/19q-codeleted oligodendrogliomas as determined by criteria in the fourth edition of the WHO classification of CNS tumors. Methylation-based molecular

subtype labels from version 11 of the random forest classifier released by Capper et al. were also downloaded. Patient age (N=420) was determined from published clinical variables, but outcome data was not available.

REMBRANDT Prediction Glioma Dataset

Binary CN4.cnchp files from Affymetrix Human Mapping 50K Hind240 (N=240) and 50K Xba240 SNP arrays (N=192) for 275 samples were downloaded from the REMBRANDT Database (GEO Data Set GSE108475) [76]. Affymetrix Power Tools (February 2021, date last accessed) was used to convert the CN4.cnchp files into text files. Precomputed copy number and loss of heterozygosity analysis results from the CN4 algorithm were extracted from these files, and HmmMedianLog2Ratio values were used to estimate the underlying DNA copy number variation using the Bioconductor package DNAcopy [77]. GISTIC was then applied to calculate gene-level gains and losses. In general, we found that data produced by Hind SNP arrays were cleaner than those generated by Xba SNP arrays and thus we used Hind-derived SCNA data for patients who had both Hind and Xba data (N=157). Clinical variables including overall survival (N=220) and age (N=208) were also downloaded. Patient age ranges (i.e., 70-74) were replaced by their median age (i.e., 72).

GISTIC 2.0 Parameters

GISTIC 2.0 (GISTIC) analysis was computed in the same manner across all datasets we generated from segmentation files [68]. GISTIC was run with the following parameters: Amplification Threshold=0.1; Deletion Threshold=0.1; Cap Values=1.5; Broad Length Cut-off=0.7; Remove X Chromosome=0; Confidence Level=0.99; Join Segment Size=4; Arm Level Peel Off=1; Maximum Sample Segments=2,000; Gene GISTIC=1; Q-value Threshold=0.25; Savegene=1; Run Broad Analysis=1; Collapse Method=extreme. The GISTIC default hg19 reference was used for hg19 alignment; the hg38.UCSC.add_miR.160920.refgene.mat file was used for alignment to hg38. All SCNA datasets other than a version of TCGA SCNA data downloaded from the GDC Data Commons (GDC hg38) were aligned to hg19.

Machine Learning Methods

We first formatted our SCNA data so that our model was robust to data acquired from older cytogenetic array technologies as well as variations in data processing pipelines. Downsampling SCNA data to chromosome arm-level resolution provided an effective solution, because such processing created a data representation that did not depend on precise gene location or GISTIC output dimension. To downsample SCNA data processed by GISTIC, we first considered amplifications (GISTIC score 2) as gains (GISTIC score 1) and homozygous deletions (GISTIC score -2) as single copy deletions (GISTIC score -1) so that all scores were between -1 and 1. Because all data is in the same range, we did not perform additional normalization. Next, we assigned each chromosome arm the mean value of the set of GISTIC scores that corresponded to that chromosome arm's genes. We ignored chromosomes X and Y as well as chromosome arms 13p, 14p, 15p, 21p, and 22p, because they had a low gene count and therefore were sensitive to noise. The final model input format was a 39-dimensional chromosome arm-level SCNA data representation. To justify this data representation, we also compared it to a 50-dimensional PCA-reduced representation of gene-level SCNA data and a 50-dimensional PCA-reduced representation of averaged cytoband-level input.

Our system consists of two stages: the first phase filters out oligodendrogliomas by screening for 1p/19q-codeletions, and the second phase passes the remaining diffuse astrocytic tumors though a binary IDH-mutation classifier. We trained and evaluated a host of machine learning classifiers that predicted IDH-mutations in adult diffuse astrocytic glioma. These classifiers were implemented using the Python packages scikit-learn. We used the Python package PyCaret [143] to prototype, tune, and calibrate our models. Area under the receiver operating characteristic curve (AUC) was maximized during model tuning. Random hyperparameter searches were conducted via 10-fold cross-validation within the cross-validation loop for cross-validated results, and on the entire training set for predictions on the four held-out validation sets. We report results from L2-penalized logistic regression, random forest, multilayer perception, and support vector machine with a radial kernel, as well as

an ensemble of these models. All cross-validated results are reported as the average of 1000 cross-validation trials. All hyperparameter choices, including model class, were made during cross-validation and only one model was applied to the independent validation datasets and the holdout TCGA validation set. All UMAP (Uniform Manifold Approximation and Projection) [144] embeddings were generated from gene-level GISTIC scores using 15 nearest neighbors and the Manhattan distance metric.

Interpretability

We used SHapley Additive exPlanations (SHAP), as implemented in the Python package shap [145], to interpret our patient-level model predictions. The SHAP algorithm assigns a value to each feature used to represent a patient that indicates how responsible that feature is to the prediction the model gives to that specific patient. The SHAP algorithm also considers all features at once, rather than independently, and thus captures feature interactions. To get patient-specific SHAP values for cross-validated results, we tracked each model that correctly or incorrectly classifies each sample over 1000 cross-validation trials. Furthermore, because each model was calibrated, we had to average the SHAP value for each uncalibrated base estimator for each of 10 calibration folds per model before averaging these values across all cross-validation trials in which the sample was correctly classified and all of those in which the sample was incorrectly classified. This process led to two SHAP values for each feature for each patient: the average SHAP value for the cross-validation trials during which the patient was misclassified and the average SHAP value for cross-validation trials during which the patient was correctly classified.

Statistical analysis

We evaluated the performance of our models with the following metrics: area under the receiver operating characteristic curve (AUC), balanced accuracy (bal accuracy), F1 score, precision, recall, and Mathews correlation coefficient (MCC). Their definitions are given below in terms of true positive (TP), true negative (TN), false positive (FP), and false

negative (FN) predictions.

$$\text{F1} = \frac{TP}{TP + \frac{1}{2}(FP + FN)}$$
$$\text{Precision} = \frac{TP}{TP + FP}$$
$$\text{Recall} = \frac{TP}{TP + FN}$$
$$\text{Balanced Accuracy} = \frac{1}{2}\left(\frac{TP}{TP + FN} + \frac{TN}{TN + FP}\right)$$

We focused primarily on the AUC and MCC scores. We preferred MCC over F1, accuracy, precision, and recall because MCC does not depend on which class we designate as positive and because MCC considers all four prediction categories (TP, FP, TN, FN) and their magnitudes [146]. We optimized for high AUC rather than MCC during model training, because AUC gives a better measure of model robustness.

6.3 Results

Overview of Global SCNAs in Adult Diffuse Gliomas

SCNA data is a promising predictor of adult diffuse glioma molecular subtype, because each adult diffuse glioma subtype exhibits a distinctive DNA structure. Oligodendrogliomas are characterized by the presence of an IDH mutation and a 1p/19q-codeletion (Figure 6.1A). Adult-type IDH-wildtype diffuse gliomas predominantly display simultaneous gain of whole chromosome 7 and loss of whole chromosome 10 (+7/-10) (Figure 6.1B). IDH-mutant astrocytomas have comparatively fewer large-scale SCNAs (Figure 6.1C). Furthermore, unsupervised methods have shown that these subtypes are largely separable by SCNA data even in low dimensions (Figure 6.1D) [95, 15, 46]. This indicates that a supervised system can robustly predict patient adult diffuse glioma molecular subtype from tumor SCNA data.

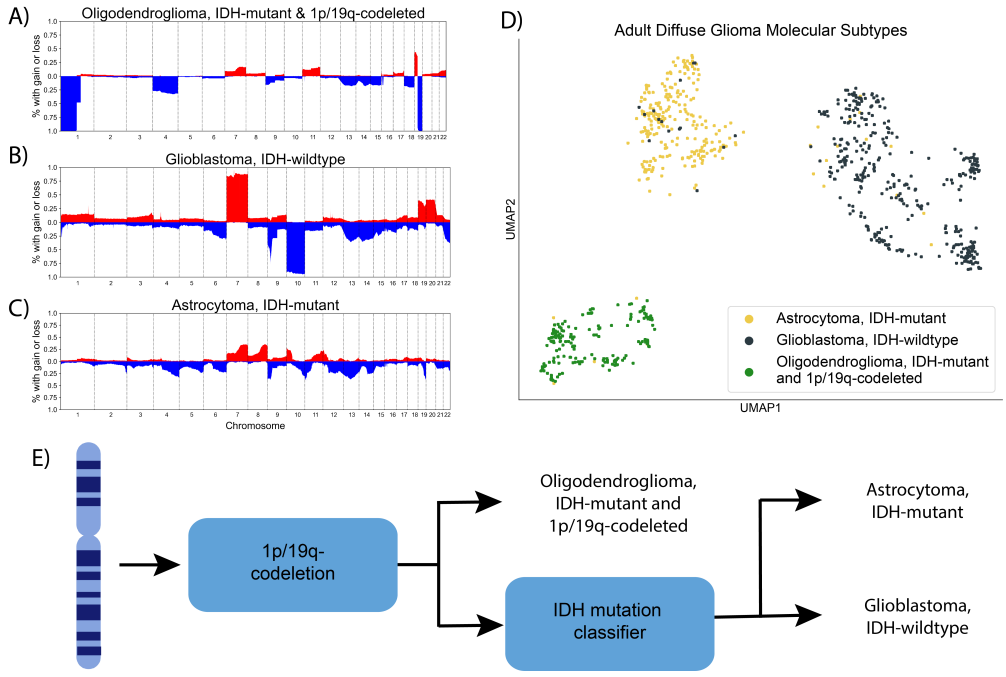


Figure 6.1: Somatic copy number alteration (SCNA) profiles of adult diffuse glioma molecular subtypes in the TCGA. A, B, C) Copy number frequency plots showing SCNA profiles for each of the three dominate molecular subtypes of adult diffuse glioma. D) UMAP landscape of SCNA data highlighting the separation of the major molecular subtypes based solely on global copy number information. E) Overview schematic of our approach to use SCNA profiles to predict IDH status in diffuse gliomas.

SCNA-based models can also incorporate domain knowledge regarding glioma subtype DNA structure, which would otherwise be difficult to learn in a purely data-driven manner. All oligodendrogliomas harbor translocation-mediated 1p/19q-codeletions that result in single-copy loss of chromosome arms 1p and 19q and intact status of chromosome arms 1q and 19p [147]. Because 1p/19q-codeletions are directly computable from SCNA data, we proposed a two-stage classification system for the prediction of adult diffuse glioma molecular subtype in which 1p/19q-codeleted oligodendrogliomas are identified in the first phase and the remaining diffuse astrocytic tumors are passed through a binary IDH mutation clas-

sifier in the second phase (Figure 6.1D). Additionally, we computed +7/-10 to verify that all tumors in our TCGA training were adult diffuse gliomas. Histological lower-grade IDH-wildtype diffuse gliomas without confirmed +7/-10, *EGFR* amplification, or *TERT* promoter (*TERT*p) mutation are considered to be pediatric-type diffuse gliomas and are genetically and biologically distinct from adult diffuse gliomas, and thus less relevant to our current study [12, 6, 19].

Determining Thresholds for 1p/19q-codeletion

Determining 1p/19q-codeletions from gene-level SCNA data requires setting a threshold for the proportion of genes lost on chromosome arms 1p and 19q and the proportion of genes retained on chromosome arms 1q and 19p. To establish such a threshold, we considered TCGA SCNA data derived from three different pipelines (UCSC hg19, GDC hg19, GDC hg38) to account for variation in data processing. For each pipeline, we observed that an 85% threshold for gene loss on chromosome arms 1p, 1q, 19p, and 19q separated all oligodendrogliomas (N=171) from all astrocytic tumors (N=615) in our TCGA training set, including several IDH-mutant astrocytomas that would have been considered 1p/19q-codeleted oligodendrogliomas using slightly lower thresholds (Figure 6.2A, Supplemental Figure A.2A). This 85% threshold was also optimal or nearly optimal for two independent validation sets containing at least 50 1p/19q-codeleted oligodendrogliomas published by Capper et al. [47] (MCC=0.97) and Jonsson et al. [73] (MCC=0.97), respectively (Figure 6.2B). Between these two validation sets, only one (0.1%) astrocytic tumor was predicted to be 1p/19q-codeleted (Supplemental Figure A.2B, C). Although six (4.6%) labeled oligodendrogliomas were predicted to be astrocytic tumors, four of the six misclassified oligodendrogliomas harbored monosomy of chromosome 1 and/or monosomy of chromosome 19, which is inconsistent with the unbalanced translocation mechanism associated with the development of oligodendrogliomas (Supplemental Figure A.2B, D) [148, 147]. Therefore, these were not model errors. The two other misclassified oligodendrogliomas fail to meet our 85% threshold, because one sample lost only 75% of 1p and the other only 71% of 19q. Regardless, we did not lower

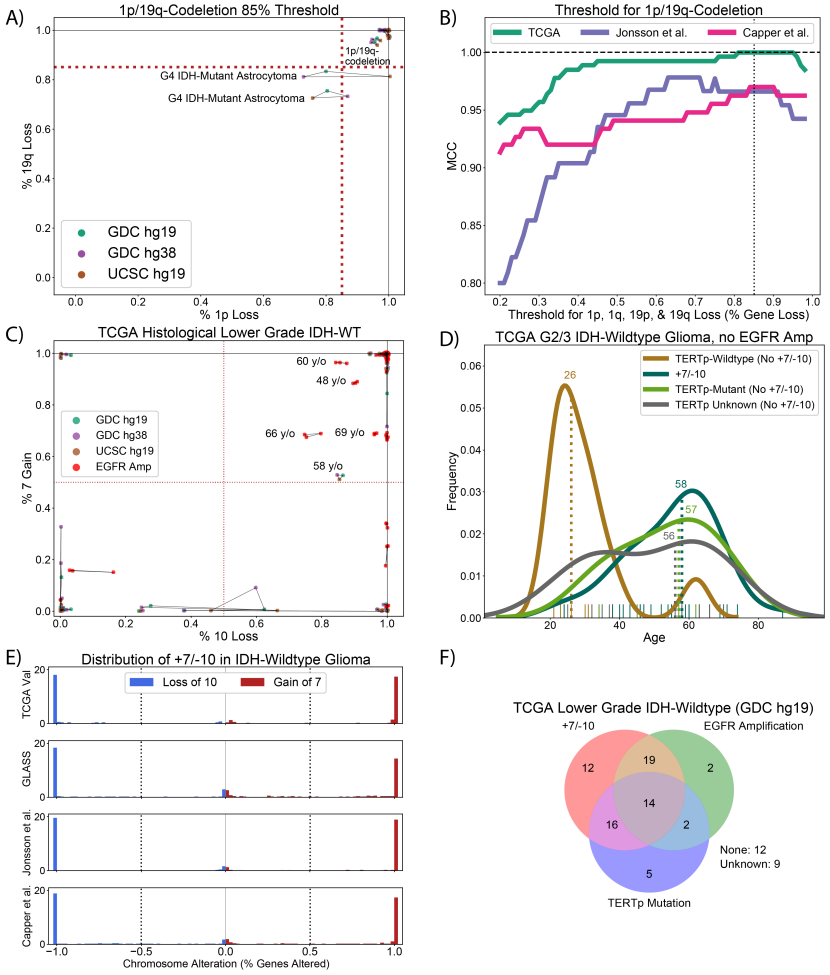


Figure 6.2: Determining rational thresholds for establishing 1p/19q-codeletion and +7/-10 status. A) Plot showing the percentage of genes lost on chromosome arms 1p and 19q for tumors showing at least 50% gene loss on 1p and 19q and at most 85% genes loss on 1q and 19p. Lines connect points that represent the same patient. B) An 85% threshold for 1p/19q-codeletions is optimal for the TCGA training set and an independent validation dataset published by Capper et al. as measured by MCC and nearly so for a dataset published by Jonsson et al. C) Plot of percent gain/loss of genes on chromosomes 7 and 10, respectively. D) Histological lower grade TCGA IDH-wildtype gliomas without *EGFR* amplification and +7/-10 are much younger than those who do. E) A 50% threshold for +7/-10 separates +7 and -10 from intact 7 and intact 10 across three independent validation sets and the TCGA holdout validation set. F) Of all TCGA histological lower-grade IDH-wildtype tumors, 21 do not or cannot be confirmed to meet the WHO 2021 criteria for adult diffuse glioma.

our threshold to include these patients, because lower thresholds risked misclassifying more IDH-mutant astrocytomas as 1p/19q-codeleted oligodendrogliomas. To assess the robustness of our 1p/19q-codeletion screen to genome coverage, this screen was applied to ULP-WGS derived TCGA SCNA data available for patients in the TCGA training set (556 astrocytic tumors, 169 oligodendrogliomas). Only one tumor was misclassified (MCC=0.996), suggesting that even ultra-low coverage (0.1x) is sufficient for accurate identification of 1p/19q-codeletions. All 1p/19q-codeletion predictions are given in the supplemental material [7].

Determining Thresholds for +7/-10

Knowledge of +7/-10 status was required to verify that all histologically-defined WHO grades 2 and 3 IDH-wildtype gliomas in our TCGA training set were adult-type diffuse gliomas [6, 19]. To ascertain +7/-10 status, we determined a threshold for the proportion of genes necessary to determine aneuploidy. This was done by identifying all TCGA patients whose adult glioma verification depended on their +7/-10 status and settling on the highest threshold for +7/-10 that did not exclude tumors whose age and outcome were consistent with quintessential IDH-wildtype glioblastoma. A 50% threshold for +7/-10 was the highest threshold capable of calling a 58-year-old short-term survivor (OS=13 months) without *EGFR* amplification an adult diffuse glioma (Supplemental Figure A.3A). Given that the next lowest threshold that would have changed a TCGA patient's adult glioma status was below 10%, we adopted 50% as our threshold for determining +7/-10 status (Figure 6.2C). As desired, this threshold divided histological lower-grade IDH-wildtype gliomas without *EGFR* amplification into two groups with significantly different age distributions (median age 35 vs. 59.5, $p < 0.0005$, Mann-Whitney U-test), especially when *TERT*p mutation status was accounted for (median age 26) (Figure 6.2D, Supplemental Figure A.3B, C). Furthermore, a 50% threshold separated the bimodal distributions of +7 and -10 for IDH-wildtype diffuse gliomas of all histological grades in three independent validation sets in addition to the TCGA validation set (Figure 6.2E). Finally, applying this 50% threshold to our TCGA training data, we excluded 21 TCGA patients: 12 lacked +7/-10, *EGFR* amplification, and

*TERT*p mutation and 9 lacked *EGFR* amplification and +7/-10 and had unknown *TERT*p mutation status (Figure 6.2F, Supplemental Figure A.3D).

IDH Mutation Classifier Design

Justification of Design Decisions

The 1p/19q-codeletion screen that comprises the first phase of our adult diffuse glioma molecular subtype predictive system passes predicted adult astrocytic gliomas to the system's second phase IDH mutation classifier. To build this classifier, we trained a logistic regression (LR) model, optimized to maximize AUC, on hg19-aligned SCNA data (GDC hg19) downsampled to chromosome arm-level resolution. The choices made during model development were justified by cross-validated experiments on the 615 adult astrocytic gliomas in our TCGA training set. Downsampled chromosome arm-level SCNA data performed as well or better than other low-dimensional SCNA input representations including PCA-reduced SCNA data (Figure 6.3A) and were more interpretable and similarly robust to noise. Indeed, downsampling to chromosome arm-level resolution had a pronounced smoothing effect on SCNA data obtained from older cytogenetic array technologies, such as the REMBRANDT study SCNA data (Figure 6.3B) [76]. We trained our model to maximize AUC performance during its parameter search and used the GDC hg19 version of the TCGA SCNA data for our training samples, because this combination of metric optimizer and dataset yielded the best cross-validated results (Figure 6.3C). We selected logistic regression as our model class, because our LR model outperformed (AUC= 0.990 +/- 0.001, MCC=0.935 +/- 0.006) a suit of other machine learning classifiers across nearly all evaluation metrics (Figure 6.3D). All predictions are given in the supplemental material [7].

Model Calibration Facilitates the Rejection of Inaccurate Predictions

To recognize patients whose model prediction may not be reliable, such as patients whose tumor is not an adult-type diffuse glioma, we calibrated our model's output probabilities and gave the option to reject low-confidence predictions. Such calibration forced the model's

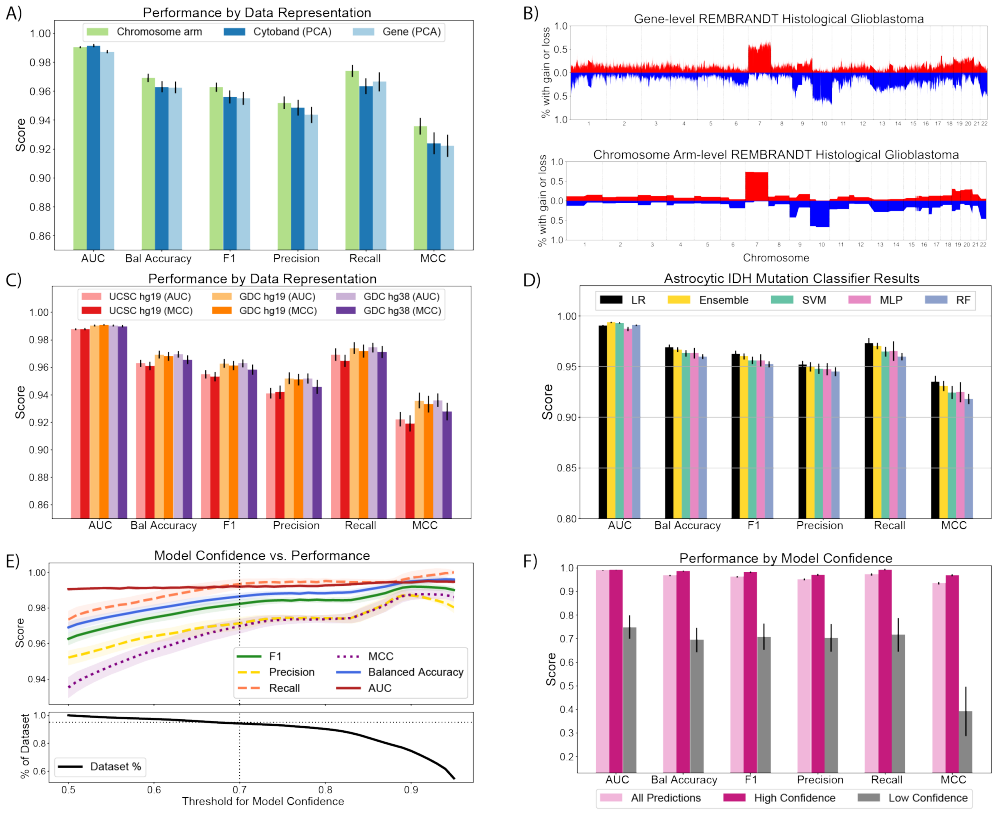


Figure 6.3: Cross-validated IDH mutation classifier development. A) Cross-validated results showed that IDH mutation logistic regression classifiers trained on chromosome arm resolution SCNA data performed better than classifiers trained on PCA-reduced gene-level or PCA-reduced cytoband-level SCNA data. B) Downsampling gene-level SCNA data to chromosome arm resolution smoothed noisy SCNA data derived from older cytogenetic arrays. C) A logistic regression model trained on TCGA SCNA data aligned to hg19 and optimized for maximizing the AUC score performed better than other parameter choices. D) Logistic regression mostly outperformed other model classes, including an ensemble of all listed models, across five metrics. E) Our model performance increased monotonically when restricted to samples of increasing prediction confidence. This indicated that the calibration of our model's output probabilities was effective. Standard deviation values for each metric over 1000 cross-validation trials are shaded in. F) Restricting model predictions to those made with confidence greater than 0.7 greatly increased model performance.

output probabilities to better represent prediction confidence, which should be low when the model evaluates tumors that do not resemble tumors in its training set. As desired, we saw that our model performed better across all metrics on patients whose prediction confidence was above 70% (AUC= 0.992 +/- 0.001, MCC = 0.970 +/- 0.004) compared to patients with lower confidence predictions (AUC= 0.75 +/- 0.05, MCC = 0.39 +/- 0.1), despite excluding only 5% of the dataset (Figure 6.3E, F).

Interpretation of IDH Mutation Classifier Cross-Validation Predictions

Misclassified Patients were Rare and Involved +7/-10

Given the robustness of our IDH mutation classifier, we were particularly interested in the rare patients it misclassified. Across 1000 cross-validation trials, 3.4% (N=12) of IDH-wildtype glioblastomas and 2.3% (N=6) of IDH-mutant astrocytomas were misclassified in at least 50% of trials (Figure 6.4A). Misclassified IDH-mutant astrocytomas had increased copy number burden, especially on chromosome 10, were disproportionately WHO grade 4 ($p<0.0001$, Fisher's Exact), and followed a clinical course significantly worse than correctly classified IDH-mutant astrocytomas (OS=2.8 vs 7.3 years, HR = 1.97, $p<0.001$, log-rank) (Supplemental Figure A.4A, B, C). Most misclassified IDH-mutant astrocytomas were embedded in a cluster of IDH-wildtype tumors defined by +7/-10 on a UMAP SCNA landscape, indicating that the DNA structure of these tumors resembles that of IDH-wildtype glioblastomas (Figure 6.4B). On the other hand, misclassified IDH-wildtype glioblastomas tended to have fewer SCNAs than their correctly classified counterparts, especially on chromosomes 7 and 10, and primarily inhabited a region on the UMAP landscape occupied by IDH-mutant astrocytomas (Supplemental Figure A.4D). Unlike IDH-mutant astrocytomas, however, no difference in histological grade or patient outcome between correctly and incorrectly classified IDH-wildtype glioblastomas was observed (Supplemental Figure A.4E, F).

Chromosome Arm 10q Holds Greatest Influence Over IDH Mutation Classification

The SHapley Additive exPlanations (SHAP) algorithm assigns a value to each chromosome

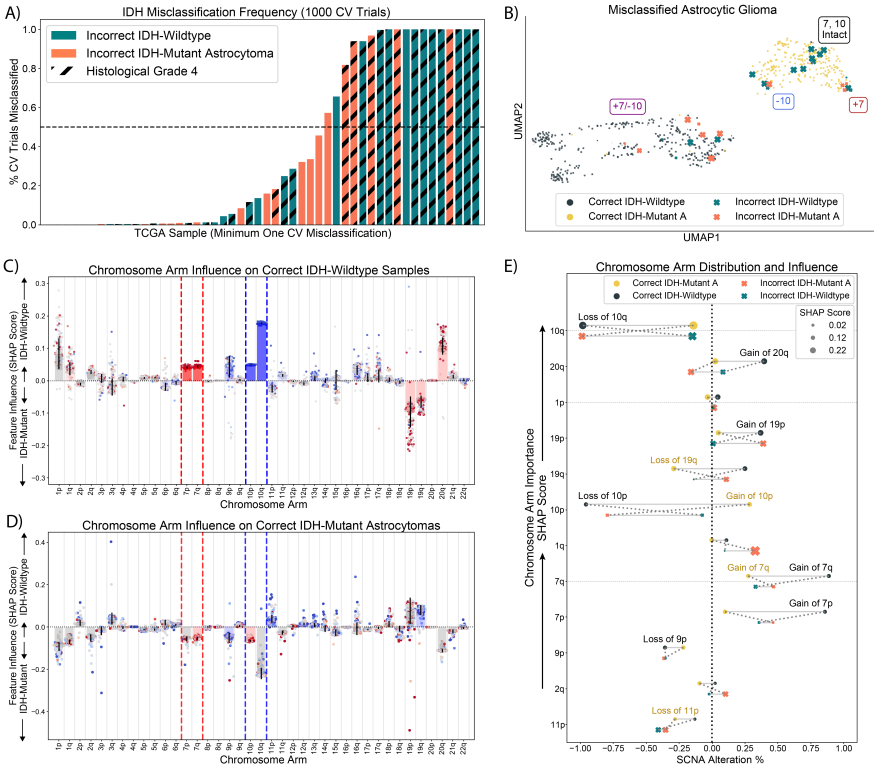


Figure 6.4: Interpretation of cross-validated results. A) Only 3.4% of IDH-wildtype gliomas and 2.3% of IDH-mutant astrocytomas are misclassified in over 50% of trials. B) Misclassified samples show regionality on a UMAP SCNA landscape. Point-size of misclassified samples indicates the frequency of misclassification. C) SHAP results of correctly classified IDH-wildtype tumors indicate that chromosome arm 10q drove correctly classified IDH-wildtype predictions. Positive vertical values favor IDH-wildtype predictions; negative values indicate a preference for IDH-mutant astrocytoma predictions. Blue indicates loss, and red indicates gain. The magnitude of the bars drawn for each chromosome arm is the average chromosome arm SHAP score for all patients plotted. D) SHAP results of correctly classified IDH-mutant astrocytomas; intact 10q drives these predictions. E) A visualization of average SHAP score and average chromosome arm SCNA status over all correct and incorrect model predictions of IDH-wildtype and IDH-mutant astrocytomas shows the top 12 chromosome arms' influence on model predictions. Solid lines connect correct or incorrect predictions; dotted lines connect groups of the same IDH subtype. Crossing dotted lines identify chromosome arms whose SCNA profile mimics the opposite IDH subtype when mistaken by our model.

arm for each patient that indicates how responsible that chromosome arm's SCNA status is to the prediction the model gives that patient [145]. A SHAP analysis of our IDH mutation classifier's cross-validated predictions indicated that loss of chromosome arm 10q drove IDH-wildtype predictions and that gain of chromosome arms 20q and 1p were more discriminative features than 10p loss or 7p and 7q gain (Figure 6.4C). Similarly, intact or marginal loss of 10q, intact 20q, intact chromosome 1, and marginal gains of 10p drove correct IDH-mutant astrocytoma predictions (Figure 6.4D). Unsurprisingly, intact 10q in IDH-wildtype gliomas or loss of 10q in IDH-mutant astrocytomas were the primary drivers of astrocytic tumor misclassification (Supplemental Figure A.5). In general, our model's mistakes were intuitive: a small subset of tumors from each subtype exhibited SCNA patterns more consistent with the opposite subtype on chromosome arms weighed heavily by the model, especially chromosome arm 10q (Figure 6.4E).

Validation of the IDH Mutation Classifier on Three Independent Datasets and the TCGA Validation Dataset

Validation Results

Our IDH mutation classifier performed well across three independent validation datasets and our holdout TCGA cohort of histological grade 4 patients with surrogate IDH labels not found from IDH sequencing. As observed during cross-validation, model performance increased with prediction confidence (Figure 6.5A, Supplemental Figure A.6). When evaluated on patients with model confidence greater than 70%, our IDH mutation classifier achieved AUC scores greater than 0.95 on each dataset (Figure 6.5B). Recall was substantially higher than precision, indicating that the model performed better on IDH-mutant astrocytomas than IDH-wildtype gliomas (Figure 6.5C). Our IDH mutation classifier also performed well (AUC = 0.98) on a version of our TCGA validation datasets whose SNCA data was derived from ULP-WGS data, indicating that this model is robust to ultra-low genome coverage (Supplemental Figure A.7). Results for model predictions on all patients with diffuse astrocytic gliomas are given in the supplemental material (Supplemental Figure A.8) [7].

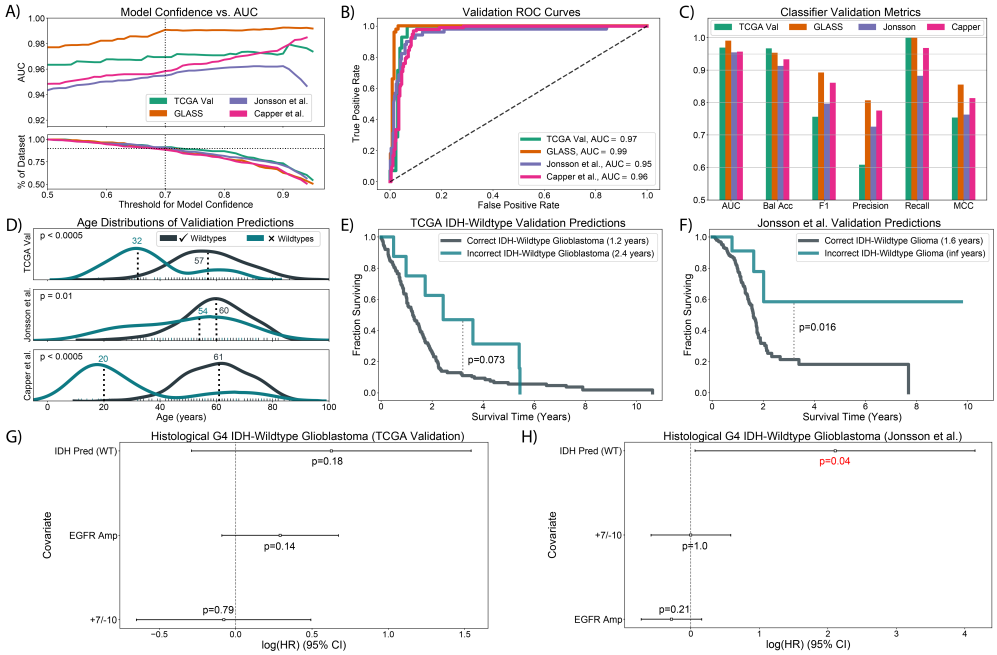


Figure 6.5: Validation results for four independent glioma datasets. A) The IDH mutation classifier's performance increased with prediction confidence on four validation sets, confirming the utility of model calibration. B) ROC curves show that our LR model performs well on four separate validation sets. C) Besides AUC, our model performs well over four additional metrics. D) Misclassified IDH-wildtype tumors are significantly younger than their correctly classified counterparts in three validation datasets (Mann-Whitney U). E, F) Misclassified IDH-wildtype tumors in the TCGA validation set and a dataset published by Jonsson et al. tend to have better outcomes than correctly classified IDH-wildtype diffuse gliomas. G, H) Plots of the results of two CPH models of histological grade 4 IDH-wildtype glioblastomas that incorporated our model's IDH mutation prediction, *EGFR* amplification, and +7/-10. Our model's IDH mutation prediction's hazard ratios are higher than those of +7/-10 and *EGFR* amplification in both validation sets.

Misclassified IDH-Wildtype Diffuse Gliomas Were Younger and Lived Longer

Misclassified IDH-wildtype tumors were significantly younger than correctly classified IDH-wildtype tumors in our TCGA validation set ($p<0.001$, Mann-Whitney U test) and validation sets published by Jonsson et al. ($p<0.01$, Mann-Whitney U test) and Capper et al. ($p<0.001$, Mann-Whitney U test) (Figure 6.5D). On two validation sets that had patient outcome data, patients with IDH-wildtype gliomas that were predicted to be IDH-mutant astrocytomas demonstrated significant ($p = 0.016$, Jonsson et al., log-rank) or marginally significant ($p=0.073$, TCGA validation, log-rank) improvement in survival compared to correctly predicted IDH-wildtype gliomas (Figures 6.5E, F). Interestingly, these results remained significant when restricted to histological WHO grade 4 tumors (Supplemental Figure A.9). No consistent age or survival difference was observed in IDH-mutant astrocytomas (Supplemental Figure A.10). Likewise, the same IDH-wildtype glioma age and outcome associations observed in the other validation sets did not hold in the GLASS, likely due to patient inclusion bias, which is inherent in datasets made up of tumors known to have second resections (Supplemental Figure A.11) [15].

Because our model's misclassified IDH-wildtype patients were associated with the absence of +7/-10 and younger age—attributes of pediatric-type tumors—we tested whether the status cIMPACT-NOW update 3 molecular markers explained the improved survival of misclassified histological WHO grade 4 IDH-wildtype tumors. The status of +7/-10 and *EGFR* amplification, but not *TERT*p mutation, were used in multivariate Cox proportional hazard models because too few *TERT*p mutation labels were available. In these analyses, the hazard ratio of the model's IDH mutation prediction was higher than either that of +7/-10 or *EGFR* amplification in the TCGA validation set and Jonsson et al. dataset, although the confidence intervals varied (Figure 6.5G,H). This evidence indicated that the observed survival benefit of misclassified histological grade 4 IDH-wildtype glioblastoma was due more to the model's prediction than +7/-10 or *EGFR* amplification status. Comple-

mentary univariate tests of combinations of +7/-10 and *EGFR* amplification in the TCGA validation set and the Jonsson et al. validation set showed inconsistent prognostic utility of these markers among histological grade 4 IDH-wildtype glioblastoma (Supplemental Figure A.12). For example, *EGFR* application was not prognostic in the Jonsson et al. dataset, and the absence of +7/-10 did not convey significant survival benefit in the TCGA validation set. With limited *TERT*p mutation status information, the clinical utility of the cIMPACT-NOW update 3 molecular markers cannot be fully assessed, but the improved survival of misclassified histological WHO grade 4 IDH-wildtype tumors is unlikely a reflection of these molecular markers as we measure them.

Many Misclassified IDH-Wildtype Glioblastomas Harbor Pediatric-Type SCNA Profiles

Of all misclassified histological WHO grade 4 IDH-wildtype glioblastoma patients in our four validation sets, only 19% showed IDH-wildtype glioblastoma-like SCNA features +7/-10 or *EGFR* amplification (Supplemental Figure A.13A). While *TERT*p mutation status was unavailable, pediatric tumors, originally diagnosed as IDH-wildtype glioblastoma, were present in the cohort of misclassified IDH-wildtype glioblastomas published by Capper et al. In this dataset, 88% of misclassified IDH-wildtype samples were diffuse hemispheric glioma, H3 G34-mutant, WHO grade 4 while only 2% of correctly classified samples were diffuse hemispheric glioma, H3 G34-mutant ($p<1e\text{-}22$, Fisher's exact, Supplemental Figure A.13B). The presence of these tumors, common in adolescent and young adults, explained the age difference between correctly and incorrectly classified IDH-wildtype glioblastomas in the dataset published by Capper et al. (Supplemental Figure A.13C) [149].

IDH and 1p/19q-Codeletion Prediction Results in the REMBRANDT Dataset

REMBRANDT Predictions and Outcomes Comparisons

Prior to applying our system to the REMBRANDT dataset, we retrained the IDH mutation classifier on all patients in our training and validation sets other than the 41 diffuse hemispheric glioma, H3 G34-mutant, WHO grade 4 provided by Capper et al. (N=1729).

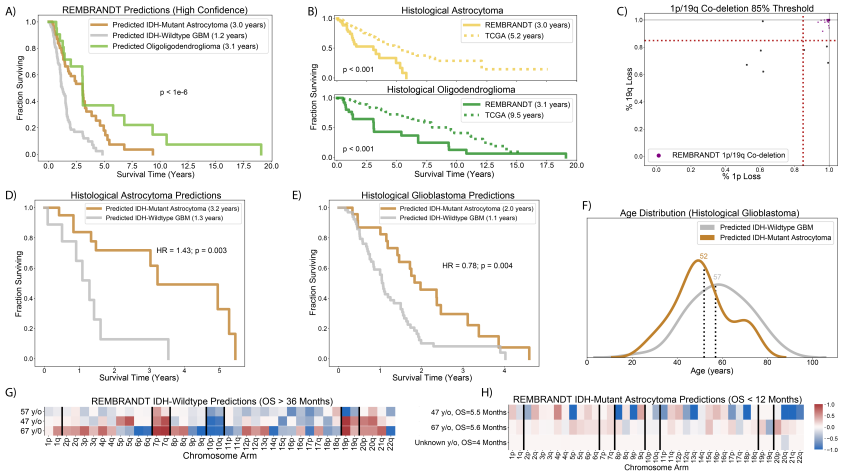


Figure 6.6: A) Predictions for REMBRANDT patients show typical median survival for IDH-wildtype glioblastoma but worse outcomes for predicted IDH-mutant astrocytomas and oligodendrogliomas compared to those in the TCGA. B) Unlike histologically-defined glioblastomas, histological astrocytomas and oligodendrogliomas in the REMBRANDT fair significantly worse than TCGA histological astrocytomas and oligodendrogliomas. C) An 85% threshold for chromosomes 1p, 1q, 19p, and 19q successfully captures the densest area of patients who pass a 50% threshold for 1p/19q-codeletion. D) Within the cohort of REMBRANDT histological astrocytomas, predicted IDH-wildtype tumors fare significantly worse than predicted IDH-mutant astrocytomas. E) Within the cohort of REMBRANDT histological glioblastomas, predicted IDH-wildtype tumors' outcomes are significantly worse than predicted IDH-mutant astrocytomas. F) Predicted IDH-wildtype tumors do not show a subpopulation of young patients that would indicate the presence of pediatric-type tumors or misclassified IDH-mutant astrocytomas. G) SCNA profiles of predicted IDH-wildtype tumors that live longer than 3 years. Three tumors show +7/-10, of which two show co-gain of chromosomes 19 and 20, and are likely IDH-wildtype tumors. H) SCNA profiles of predicted IDH-mutant astrocytoma that live less than 12 months. None show SCNA characteristics of IDH-wildtype glioblastoma.

Predictions on all REMBRANDT patients with model confidence greater than 70% show that predicted IDH-wildtype glioma, regardless of tumor histological grade, have similar survival trajectories and median overall survival (OS = 1.1 years) as IDH-wildtype glioblastomas in the TCGA (OS=1.2 years) (Figure 6.6A). On the other hand, predicted IDH-mutant astrocytomas (OS=3 years) and predicted IDH-mutant and 1p/19q-codeleted oligodendrogliomas (OS=3.1 years) in the REMBRANDT dataset have significantly shorter median survival than TCGA IDH-mutant astrocytomas (OS=7.3 years) and 1p/19q-codeleted oligodendroglioma (OS=11.2 years), respectively. This discrepancy can be attributed to differences in average outcomes of histological astrocytomas (OS=3.0 vs. 5.2 years) and histological oligodendrogliomas (OS=3.1 vs. 9.5 years) between the REMBRANDT study and the TCGA (Figure 6.6B). Predicted REMBRANDT IDH-mutant astrocytomas are also disproportionately older and higher grade than TCGA IDH-mutant astrocytomas, although the same is not true of predicted REMBRANDT and TCGA 1p/19q-codeleted oligodendrogliomas (Supplemental Figure A.14). Results for model predictions on all REMBRANDT patients are given in the supplemental material (Supplemental Figure A.15) [7].

REMBRANDT 1p/19q-Codeletion Screen Highlights Limitations of Histological Only Diagnosis

Our REMBRANDT 1p/19q-codeleted oligodendroglioma screen captured the densest area of tumors with significant gene losses on chromosome arms 1p and 19q, consistent with the 1p/19q-codeletion screens on our training and validation sets (Figure 6.6C). As reported elsewhere, our predictions highlight the difficulty of absolutely distinguishing astrocytomas from oligodendrogliomas based on histomorphology alone [150, 151]. Only 46% of REMBRANDT histological oligodendrogliomas harbor predicted 1p/19q-codeletions, and only 48% of REMBRANDT tumors harboring predicted 1p/19q-codeletions were diagnosed as histological oligodendroglioma (Supplemental Figure A.16A). Visual inspection of patient copy number profiles showed clear evidence of 1p/19q-codeletions in our predicted oligodendroglioma patients (Supplemental Figure A.16B). Similarly, in patients diagnosed with oligodendroglioma

for whom we did not find 1p/19q-codeletions, we saw that most (5/6) tumors that lose 1p harbor monosomy chromosome 19, prohibiting a 1p/19q-codeletion (Supplemental Figure A.16C). The exception was a single patient with full loss of 1p and intact 1q and 19p, but whose proportion of 19q loss falls slightly below our 85% threshold (79%): this patient's tumor was likely an oligodendroglioma (Supplemental Figure A.16D).

REMBRANDT IDH Mutation Classifier Results on Histological Astrocytic Tumors

Following our screen for oligodendrogliomas, we applied our IDH-mutation classifier to patients diagnosed with histological astrocytoma and glioblastoma not harboring predicted 1p/19q-codeletions. Our IDH mutation predictions on histological astrocytomas generated a dramatic survival difference between predicted IDH-wildtype gliomas and IDH-mutant astrocytomas (HR=1.43, $p = 0.003$, log-rank) and appeared to correctly identify histological lower-grade IDH-wildtype gliomas (OS=1.1 years) now considered to be IDH-wildtype glioblastomas (Figure 6.6D). Histological glioblastomas predicted to be IDH-wildtype glioblastomas had the same median overall survival (1.1 years), and their survival trajectory was significantly worse than those of histological grade 4 tumors that were predicted to be IDH-mutant astrocytomas (HR=0.78, p=0.004, log-rank) (Figure 6.6E).

Possible IDH Mutation Classifier REMBRANDT Errors

To identify possible IDH mutation classification errors on the REMBRANDT dataset, we examined abnormally young or long-living predicted IDH-wildtype glioblastomas, as well as abnormally old or short-living, predicted IDH-mutant astrocytomas. We did not observe a subset of significantly younger patients within the cohort of predicted IDH-wildtype patients which may have represented misclassified IDH-mutant astrocytomas (Figure 6.6F). Of the four predicted IDH-wildtype glioblastoma patients that lived longer than three years, two harbored hallmark +7/-10 and *EGFR* amplification as well as co-gain of chromosomes 19 and 20, a documented marker for long-term survivors in IDH-wildtype glioblastoma [99] (Figure 6.6G). These tumors were likely correctly classified. The remaining tumors were inconclusive.

One lacked +7/-10 but its patient's age (57 years old) was consistent with IDH-wildtype glioblastoma, and the other harbored +7/-10 but was atypically young (32 years old). Both lacked *EGFR* amplification. Among predicted IDH-mutant astrocytomas, we observed an older subset of patients (age range 65-80, N=4). However, these tumors had few IDH-wildtype glioblastoma-like SCNA features: none had +7/-10 and only one displayed *EGFR* amplification (Supplemental Figure A.17). Similarly, of the six IDH-mutant astrocytomas that lived less than 12 months, none harbored +7/-10 or *EGFR* amplification, and two had nearly zero copy number alterations (Figure 6.6H). These tumors were likely IDH-mutant astrocytomas, although the lack of *TERT*p mutation knowledge may have hidden IDH-wildtype glioblastomas.

6.4 Discussion

We developed a system that predicts adult diffuse glioma molecular subtype from SCNA data, verified its accuracy on three independent validation datasets in addition to the TCGA validation dataset, and applied it to the retrospective REMBRANDT study. In a platform-independent manner, this system can robustly assign molecular subtype labels to patients with SCNA data derived from several molecular methods, including SNP-array, methylation array, whole-exome sequencing, and whole-genome sequencing. This is relevant because retrospective glioma studies with SCNA data, but no molecular subtype information, can be transformed into effective validation datasets for contemporary research. Furthermore, the gene-level thresholds we proposed for calling 1p/19q-codeletions and +7/-10 are applicable to all glioma datasets with SCNA data, regardless of whether IDH information is available.

The 85% threshold we used for the 1p/19q-codeletion screen in our system's first phase leveraged data from 786 gliomas and was validated on data from 940 tumors across two datasets. Despite the abundance of evidence supporting this threshold, we recommend visually inspecting the SCNA profiles of patients who fall near this threshold. We suspect that there exist IDH-mutant astrocytomas that lose 1p and 19q by means other than translocation and should not be considered 1p/19q-codeleted. We identified one such patient (Supplemen-

tal Figure 2C) who also showed a remarkably high SCNA burden. We conjecture that in this case 1p and 19q loss, along with other large-scale SCNAs, may be the result of stochastic processes. Accordingly, tumors harboring high SCNA burden that are predicted to harbor 1p/19q-codeletion by our system should be flagged for closer examination. Conversely, there are rare 1p/19q-codeleted oligodendrogliomas that do not meet our 1p/19q-codeletion threshold. Having observed two such tumors (Supplemental Figure 2E), we suggest that tumors falling just shy of our 1p/19q-codeletion threshold be inspected for oligodendroglioma-like SCNA patterns outside of chromosome arms 1p and 19q. Low tumor cellularity may explain why some oligodendroglioma gliomas appear to lose significantly less than 100%, and occasionally fewer than 85%, of genes on chromosome arms 1p and 19q. It is also conceivable that IDH-wildtype tumors may meet our 1p/19q-codeletion threshold, but we saw no evidence for this.

Our threshold for +7/-10 will benefit from further corroboration. A rigorous search for a threshold for +7/-10 requires more histological lower-grade IDH-wildtype tumors with available *TERT*p mutation status and outcome data in addition to SCNA data. Only 12 such tumors were available in the TCGA dataset. Without *TERT*p mutation status, it is not possible to define two groups of histologically-defined WHO grades 2 or 3 IDH-wildtype glioma whose status as either pediatric-type gliomas or IDH-wildtype glioblastomas depends on the threshold set for +7/-10. Therefore, it is difficult to claim a particular threshold is optimal. Instead, based on our analysis, we suggest that a 50% threshold for +7/-10 is reasonable. Additional knowledge of *TERT*p mutation status is also needed for a rigorous analysis of the prognostic value of cIMPACT-NOW update 3 criteria in histological grade 4 IDH-wildtype glioblastoma. Our analysis of +7/-10 and EGFR amplification did not consistently show a survival benefit among histological grade 4 IDH-wildtype tumors that lacked these markers, but other studies have [152]. These studies were able to identify more "triple negative" (no +7/-10, no EGFR amplification, no *TERT*p mutation), primarily because their method for +7/-10 determination was less sensitive–another reason to refine our threshold for +7/-10. Knowledge of *TERT*p mutational status is also needed for tumor

types besides IDH-wildtype glioma. For example, emerging evidence suggests that *TERT*p mutations may convey positive prognoses in IDH-mutant astrocytomas [153].

In the second phase of our system, we observed that excluding low-confidence predictions improved the classifier's performance and may have filtered out unfamiliar tumors, such as non-adult diffuse glioma likely. However, our model calibration strategy does not reject all non-adult diffuse glioma, such as many diffuse hemispheric glioma, H3 G34-mutant in the dataset published by Capper et al. In future model iterations, we will train a multiclass classifier which will help identify other tumor types in the REMBRANDT dataset. Additional improvements include using data augmentation to generate synthetic WHO grade 4 IDH-mutant astrocytoma, especially those with losses on chromosome 10, which are underrepresented in our training set and are difficult to distinguish from IDH-wildtype glioblastoma. Additionally, ensembling our IDH mutation classifier with a model trained without chromosomes arms 7p, 7q, 10p, and 10q may improve the classification of IDH-wildtype tumors without +7/-10.

We observed that misclassified histological WHO grade 4 IDH-wildtype glioblastomas tended to have better outcomes than correctly classified WHO grade 4 IDH-wildtype glioblastomas in our TCGA validation set and the dataset published by Jonsson et al. Given that these tumors were all histological WHO grade 4, the existence of pediatric-type diffuse glioma does not explain their relatively favorable clinical course. We noted that these misclassified tumors rarely exhibit SCNA features commonly associated with IDH-wildtype glioblastoma such as +7/-10 and EGFR amplification. It remains an open question whether histological WHO grade 4 IDH-wildtype glioblastoma without molecular features of IDH-wildtype glioblastoma should be considered less aggressive than their counterparts, but our results support this notion [152].

In conclusion, the primary contribution of this work is the development of a computational tool that accurately classifies the molecular subtype of patients' tumors in retrospective adult diffuse glioma studies that have available SCNA data. We identified all patients with

likely 1p/19q-codeleted oligodendrogliomas, IDH-wildtype glioblastomas, and IDH-mutant astrocytomas in the REMBRANDT study in an effort to make the REMBRANDT study a better resource for validating diffuse glioma research. We also propose evidence-based thresholds for calling 1p/19q-codeletions and +7/-10 from gene-level SCNA data.

Chapter 7

AN INDEPENDENTLY PROGNOSTIC AGE-ASSOCIATED INDICATOR FOR POOR SURVIVAL IN OLIGODENDROGLIOMA

Abstract

Background: Oligodendroglioma, IDH-mutant and 1p/19q-codeleted has highly variable outcomes that are strongly influenced by patient age. This study aimed to discover novel molecular biomarkers that may characterize an aggressive subgroup of oligodendroglioma associated with, but independently prognostic of, patient age. *Methods:* Differential gene expression analysis was conducted on TCGA RNA-seq data (N=169) to identify pathways and genes that differ significantly between older and younger oligodendrogliomas. Genes whose expression was associated with age and survival in multivariate analyses were further studied using DNA methylation (N=171). The prognostic value of their expression and methylation profiles was then compared to known genomic biomarkers and histopathologic features. RNA-seq and DNA methylation validation data were obtained from the CGGA (N=44) and a cohort published by Capper et al. (N=144), respectively. *Results: HOXD12* overexpression was associated with both patient age and survival in the TCGA (FDR<0.01, FDR=1e-5) and the CGGA (p=0.03, p<1e-3). *HOXD12* gene body hypermethylation was associated with age, tumor grade, and survival in the TCGA (p<1e-5, p<0.001, p<1e-3) and with age and tumor grade in Capper et al. (p<0.02, p=0.001). In the TCGA, *HOXD12* gene body hypermethylation and overexpression were independently prognostic of *NOTCH1* and *PIK3CA* mutations, loss of 15q, and *MYC* activation as well as mitotic figures, Ki-67 index, microvascular proliferation, and necrosis. *Conclusions: HOXD12* overexpression and gene body hypermethylation are associated with an older, atypically aggressive subgroup of

oligodendroglioma and may serve as superior prognostic biomarkers compared to previously reported genomic biomarkers and standard histopathologic features.

7.1 Introduction

Oligodendroglioma, IDH-mutant and 1p/19q-codeleted is a subtype of adult diffuse glioma, molecularly distinct from IDH-mutant astrocytoma and IDH-wildtype glioblastoma [6]. Patients diagnosed with oligodendroglioma generally have favorable outcomes compared to other adult diffuse gliomas, but individual patient outcomes are highly variable, with some patients living less than a year and others living decades [42, 43, 154, 155]. Long-term quality of life must be considered when treating patients expected to experience long-term treatment effects, such as radiation-induced cognitive decline, but the existence of aggressive oligodendroglial tumors complicates discerning which patients can safely receive less aggressive treatment. In practice, this has led to a heterogeneous and controversial collection of treatment regimens [156].

The ongoing CODEL trial [157] may settle some controversy surrounding competing chemotherapy agents PCV and temozolomide, but quality of life arguments favoring delaying or forgoing radiation therapy and/or chemotherapy for certain patients will persist [158, 159]. Thus, patient-specific risk indicators are needed to identify patients with abnormally aggressive disease. To date, only a few prognostic molecular biomarkers have been proposed (*PIK3CA* mutations [160], *NOTCH1* mutations [154, 155], *CDKN2A* homozygous deletion [56], loss of 15q [155], and MYC activation [161]) to complement known clinical factors associated with survival such as age, KPS, World Health Organization (WHO) grade, and extent of resection [46, 162]. In this study, we sought to discover and validate novel prognostic biomarkers in oligodendroglioma.

Previously, we observed that the age distribution of patients diagnosed with oligodendroglioma appeared bimodal [46]. Based on our previous studies of IDH-wildtype glioblastoma, and given that oligodendroglioma age is prognostic, we hypothesized that older oligodendroglioma patients may have gene expression signatures indicative of poor survival [96,

163]. These age-related molecular biomarkers may reflect a more aggressive subgroup of oligodendroglioma that occurs disproportionately in older patients. To formalize and interrogate this hypothesis, we investigated oligodendroglioma data across multiple cohorts, including The Cancer Genome Atlas (TCGA), the Chinese Glioma Genome Atlas (CGGA), the Glioma Longitudinal AnalySiS consortium (GLASS), and a cohort of patients published by Capper et al. to identify, validate, and characterize prognostically relevant genes [59]. We also evaluated the utility of these genes as prognostic factors independent of known histopathologic, genomic, and radiographic features associated with poor survival in oligodendroglioma.

7.2 Methods

The Cancer Genome Atlas (TCGA) RNA-Seq, DNA Methylation, and Clinical Data

Precomputed DNA methylation beta values (Illumina 450K) from 171 primary oligodendroglioma, IDH-mutant and 1p/19q-codeleted tumors in The Cancer Genome Atlas (TCGA) were downloaded from the University of California Santa Cruz (UCSC) Xena (`https://xena.ucsc.edu/`) [82]. RNA-seq gene counts for TCGA oligodendroglioma were downloaded from the National Cancer Institute's Genomic Data Commons (GDC) Data Portal (`https://gdc.cancer.gov/`) (N=170), and TCGA transcript per million data (TPM) were computed from the recount2 dataset [78] (N=169), as specified by Arora et al. [79] to be consistent with available Chinese Glioma Genome Atlas (CGGA) TPM data. Characteristics including age, WHO grade, and overall survival were also downloaded from UCSC Xena, and treatment details (postoperative_rx_tx, radiation_therapy) were downloaded from the GDC Data Portal. TCGA IDH-mutational status and 1p/19q-codeletion data were ascertained as previously described [7].

TCGA Histopathologic Features

Histopathologic feature values for Ki-67 proliferation index, necrosis, microvascular proliferation, and mitotic figures were recorded for 156 oligodendroglioma patients as specified in

Section 2.6.2.

TCGA Genomic Biomarkers

The mutational status of *NOTCH1* and *PIK3CA* and somatic copy number alteration (SCNA) status of *CDKN2A* and chromosome arm 15q were determined from TCGA somatic mutation calls and SCNA GISTIC calls as previously described [7]. Chromosome 15q loss was called if more than half of the genes on 15q were lost. *CDKN2A* homozygous deletions were called if *CDKN2A* had a GISTIC score of -2 in at least two of three versions of TCGA SCNA calls [7]. MYC activation was called as described in Kamoun et al. [161].

TCGA Radiographic Features

The Multimodal Brain Tumor Segmentation Challenge (BraTS) 2020 training and validation datasets provide pre-operative, multimodal MRI tumor scans for 27 oligodendroglioma patients in the TCGA [103, 102]. Ground-truth segmentations delineating the boundaries of tumor edema, enhancement, and necrotic/non-enhancing tissue were available for the 13 oligodendrogliomas. We used a segmentation machine learning model trained on the BraTS training set to assign segmentations to the remaining 14 oligodendrogliomas followed by manual correction [16, 57]. These segmentation maps were used to calculate the volume of each tumor compartment, which was then binarized as high or low (Supplemental Figure A.18).

TCGA Pan-Cancer Data

RNA-seq FPKM data (N=10,169), DNA methylation beta values (N=8814), and clinical variables age (N=11,030) and overall survival (N=11,015) from 11,081 unique TCGA patients diagnosed with one of 33 cancer types were downloaded from UCSC Xena. Patients were classified as young or old using the median age of incidence for each cancer type as a threshold. FPKM data was converted to TPM as in Arora et al. [79].

Chinese Glioma Genome Atlas (CGGA) Dataset

Raw RNA gene counts for 48 Chinese Glioma Genome Atlas (CGGA) [80] primary oligodendroglioma tumors were downloaded from recount2 [78] and used to compute TPM values [79]. Unsupervised dimension reduction [81] of CGGA and TCGA TPM data identified four CGGA oligodendroglioma outliers; these patients were excluded from our gene expression analysis (Supplemental Figure A.19). Clinical variables including age (N=48), tumor WHO grade (N=48), and overall survival (N=46) were deduced from published annotations [80].

Capper et al. Dataset

Illumina 450K platform DNA methylation IDAT files for 144 primary oligodendroglioma tumors made available by the authors of Capper et al. [47] were downloaded from the NCBI Gene Expression Omnibus (GEO) under accession number GSE109381. Beta values were determined using the R package minfi [69]. Patient age (N=144) and WHO tumor grade (N=80) were determined from published clinical variables. Outcome data was not available. Tumors were considered oligodendroglioma if the WHO 2016 pathological diagnosis and the classification of the version 11 Capper et al. random forest classifier were both oligodendroglioma or if the reported final pathological diagnosis (after integration with the random forest classifier result) was oligodendroglioma.

The Glioma Longitudinal AnalySiS (GLASS) Dataset

The Synapse API was used to download gene expression TPM data (N=9), DNA methylation beta values from the merged Illumina 450K and EPIC platforms (beta.merged) (N=11), as well as clinical variables age (N=43) and grade (N=38) for primary oligodendroglioma tumors in the Glioma Longitudinal AnalySiS (GLASS) (Data Release version 2021-11-15) [75, 74]. All GLASS samples that appear in the TCGA were excluded.

Jonsson et al. Dataset

Published clinical variables, including overall survival and age for 136 primary oligodendroglioma tumors were downloaded from cBioPortal [71, 72, 73].

Gene Expression Analysis

Differential gene expression analysis was conducted using the R package DESeq2 [164]. Genes shown in a volcano plot were limited to the genes that have raw gene counts data and TPM data available. When testing for gene expression prognostic implications during unbiased searches, we compared the outcomes of the top third of patients to the bottom third of patients based on their TPM values. If more than a third of patients did not express that gene, we compared the patients who expressed the gene to those who did not. *HOXD12* expression was considered high/overexpressed if *HOXD12* expression was non-zero; low *HOXD12* expression indicates no *HOXD12* expression. Gene ontology analysis and gene set enrichment analysis were performed using the R package clusterProfiler [165] (R version 4.1.1).

HOXD12-Specific Methylation Analysis

We used public annotations to determine DNA methylation probe genomic regions (TSS -1500, TSS -200, gene body, or 3'UTR) [166]. The mean value of the beta values of the three probes (cg23130254, cg03964958, cg03371669) in *HOXD12*'s gene body were used to determine *HOXD12* hypermethylation. If the mean value exceeded the TCGA's 67th quartile (0.358), *HOXD12* was considered hypermethylated.

Statistical analysis

The Shapiro-Wilk test was used to test distribution normality. The log-rank test was used to test for univariate survival association. We used Fisher's exact test (Fisher's) to assess statistical differences in binary data across two groups. Cox's proportional hazard (CPH) regression was used to compute hazard ratios (HR) and for multivariate survival analysis. Covariates whose variance inflation factor (VIF) indicated collinearity (VIF>5) were omitted from CPH regression. The R package rms was used to construct a nomogram. Statistical analyses were performed in Python version 3.8.6 or R version 4.0.5 unless otherwise specified.

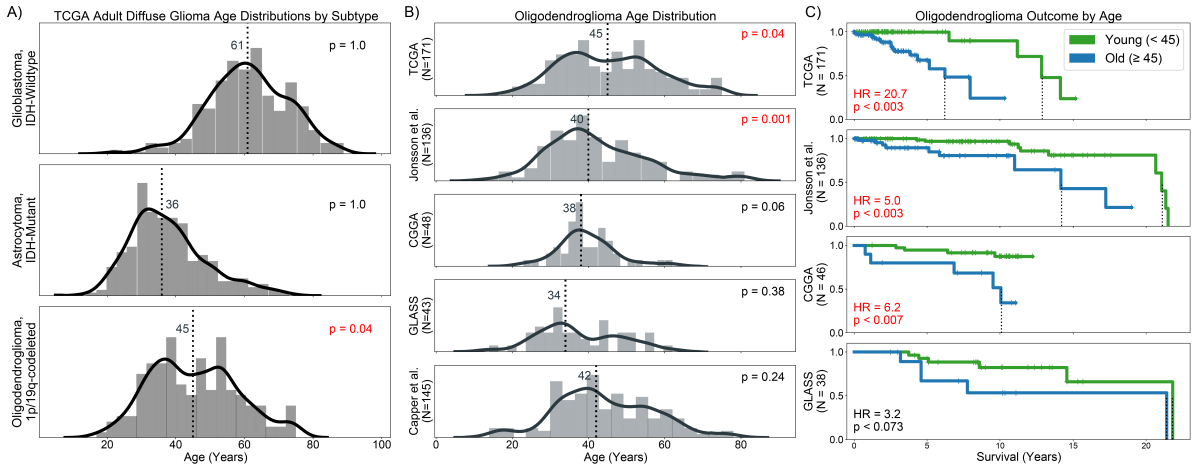


Figure 7.1: Oligodendroglioma Age is not Normally Distributed and is Related to Survival. A) Oligodendroglioma was the only subtype of adult diffuse glioma for which TCGA patient age was not normally distributed. B) The distribution of oligodendroglioma patient age was significantly and marginally significantly non-Gaussian in oligodendroglioma patients published by Jonsson et al. and the CGGA, respectively. C) Older age conveyed significantly worse outcomes in the TCGA, Jonsson et al., and CGGA cohorts.

7.3 Results

Age at Diagnosis of Oligodendroglioma is not Normally Distributed and is Related to Survival

Unlike IDH-wildtype glioblastoma and IDH-mutant astrocytoma adult diffuse gliomas in The Cancer Genome Atlas (TCGA), patient age in TCGA oligodendrogliomas was not normally distributed ($p<0.04$, Shapiro-Wilk) (Figure 7.1A). We also observed significantly and marginally significantly non-Gaussian oligodendroglioma age distributions in a cohort of 136 patients published by Jonsson et al. ($p=0.001$, Shapiro-Wilk) and in a cohort of 48 patients from the Chinese Genome Glioma Atlas (CGGA) ($p=0.06$, Shapiro-Wilk) (Figure 7.1B). Although age was not significantly non-Gaussian in cohorts published by the Glioma Longitudinal AnalySiS Consortium (GLASS) and Capper et al., histograms show visual evidence

suggestive of bimodal distributions.

In addition to validating that the oligodendroglioma age distribution was non-Gaussian, we verified that patient age was prognostic in our datasets. Splitting patients by the median age of TCGA oligodendrogliomas (45 years), older age conveyed significantly worse outcomes in the TCGA, Jonsson et al., and CGGA cohorts ($p<0.003$, $p<0.003$, $p<0.007$, log-rank) (Figure 7.1C). While an age threshold of 45 did not stratify survival among GLASS oligodendrogliomas, there were fewer older patients in GLASS (N=7, 28%) because the GLASS dataset, which requires patients to have second resections, is skewed for younger, favorable survivors. This bias also may have impeded the age distribution of GLASS oligodendrogliomas from being non-Gaussian. Altogether, these data motivate the hypothesis that there may exist a distinctive aggressive subgroup of oligodendroglioma that occurs at older ages.

Oligodendrogliomas in Older Patients are Associated with Increased Developmental Transcription Factors and Decreased Keratins

Gene ontology analysis comparing gene expression profiles of older ($\geq$45 years) and younger (<45 years) TCGA oligodendroglioma patients showed that the most significantly activated pathways in older patients with TCGA oligodendrogliomas were linked to developmental transcription factors and included DNA-binding transcription activator activity (padj<1e-6) and the development of the embryonic skeletal system (padj<1e-7), appendages (padj<1e-6), and limbs (padj<1e-6) (Figure 7.2A,B). Significantly suppressed pathways included cornification (padj<1e-13) and keratinization (padj<1e-10) and other pathways that involve structural proteins that control cell shape. Our results were supported by a gene set enrichment analysis, whose top hits were the suppression of cornification (padj<1e-13) and the activation of embryonic skeletal system development (padj<1e-7) (Figure 7.2C).

***HOXD12* is Upregulated in Older Patients with Oligodendroglioma and is a**

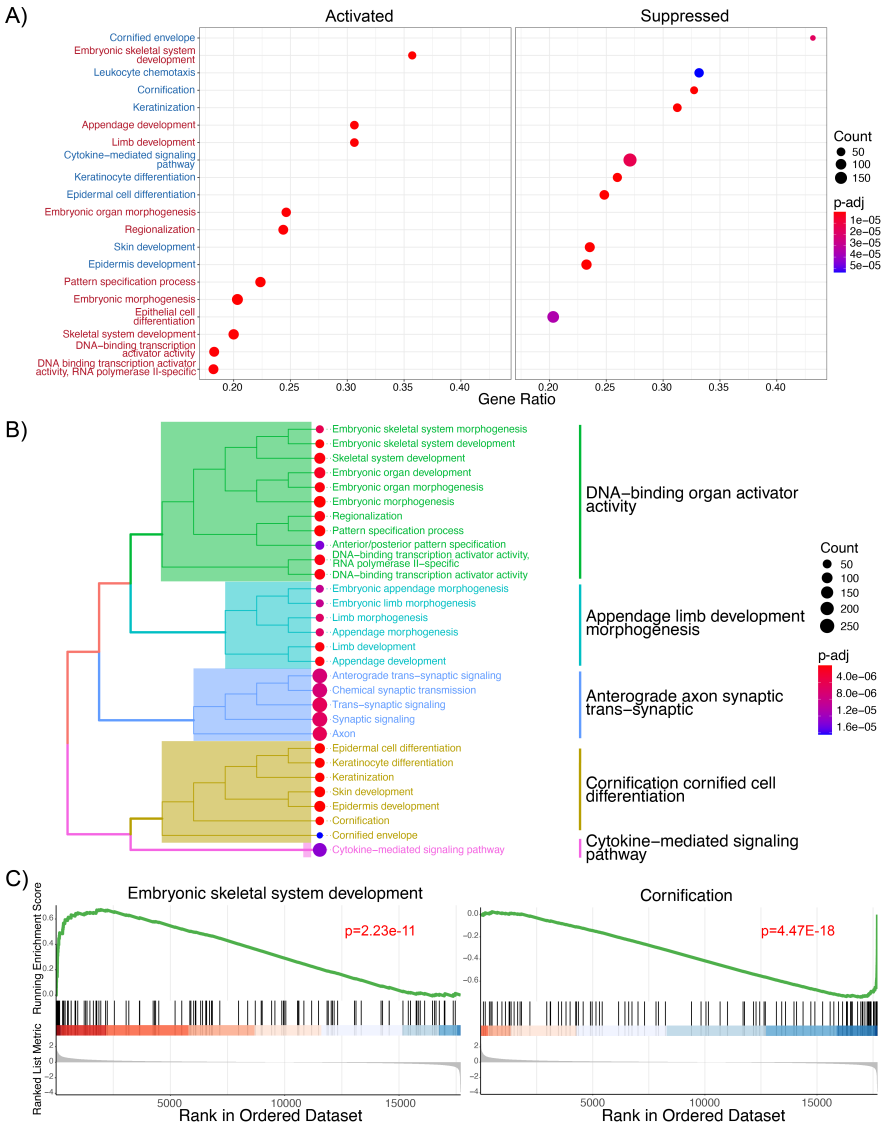


Figure 7.2: Older Oligodendrogliomas are Associated with Increased Developmental Transcription Factors and Decreased Keratins. A, B) Gene ontology analysis showed that activated pathways in TCGA oligodendrogliomas from older patients were linked to developmental transcription factors, specifically DNA-binding transcription activator activity and the development of the embryonic skeletal system, appendages, and limbs. Suppressed pathways included cornification, keratinization, and other pathways that control cell shape. C) The top hits of gene set enrichment analyses were the activation of embryonic skeletal system development and suppression of cornification.

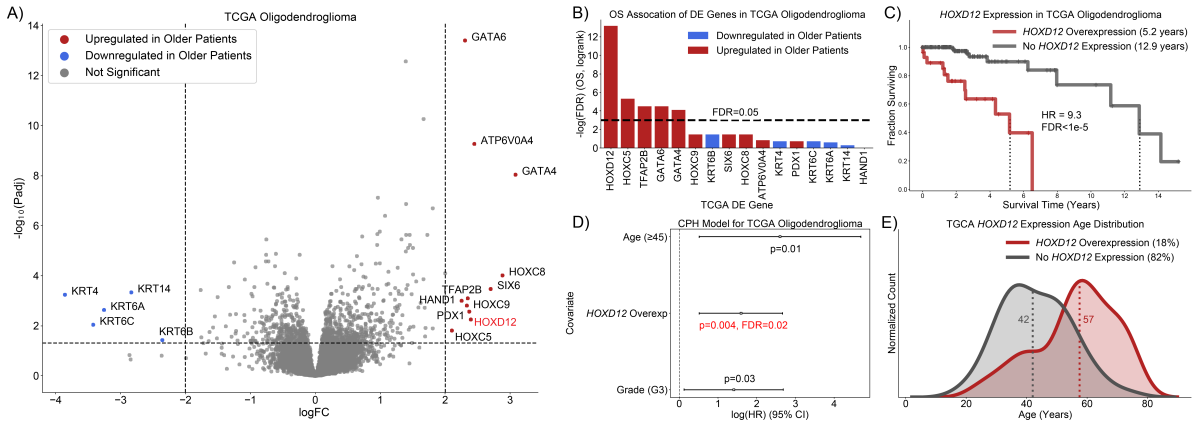


Figure 7.3: *HOXD12* is Upregulated in Older TCGA Oligodendroglioma Patients and is a Marker for Poor Survival. A) Unbiased differential gene expression analysis returned 16 significantly differentially expressed genes, including several genes in the HOX, GATA, and keratin gene families. B) Five differentially expressed genes conferred significantly worse outcomes when overexpressed. C) Kaplan-Meier curves show that oligodendrogliomas patients harboring *HOXD12* overexpression had clinical courses significantly worse than other oligodendrogliomas. D) *HOXD12* was the only gene whose overexpression was prognostic independent of patient age and WHO grade. E) The age distribution of the oligodendrogliomas without *HOXD12* overexpression was not significantly non-Gaussian.

Marker for Poor Survival

We performed survival analyses on 16 genes identified by an unbiased differential gene expression analysis that compared older and younger TCGA oligodendroglioma patients. These genes included developmental transcription factors in the HOX and GATA gene families, which were upregulated in older patients, and keratin genes, which were downregulated in older patients (Figure 7.3A). Of these 16 genes, five genes conferred significantly worse outcomes when overexpressed, of which *HOXD12* was the most significant (HR=9.3, FDR<1e-5, log-rank) (Figure 7.3B, C). Furthermore, *HOXD12* was the only gene whose overexpression was also prognostic independent of patient age and tumor grade (FDR=0.02, CPH) (Fig-

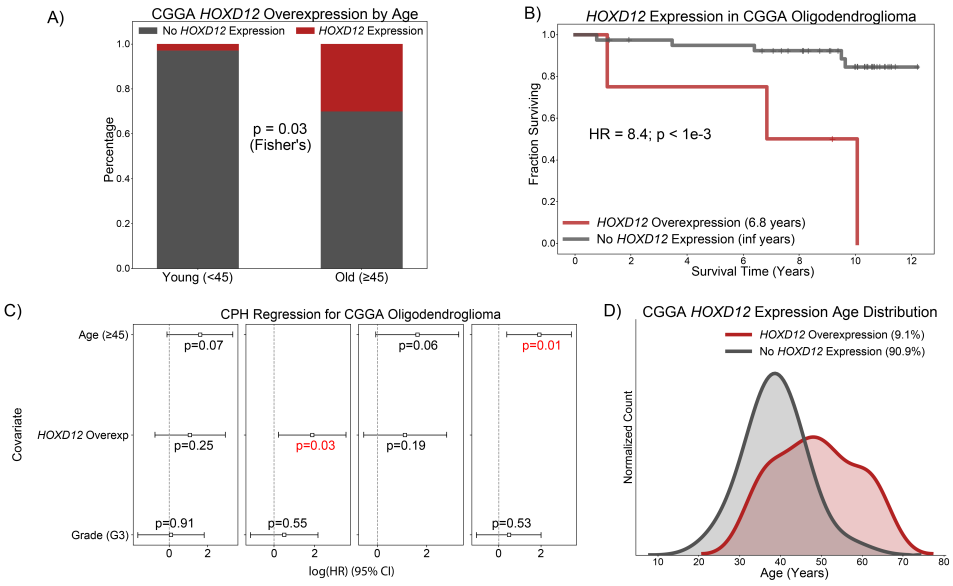


Figure 7.4: *HOXD12* Overexpression is Age-Associated and Survival-Associated in the CGGA. A) Older CGGA oligodendroglioma patients were enriched for *HOXD12* overexpression. B) HOXD12 overexpression was prognostic in the CGGA. C) *HOXD12* overexpression was prognostic independent of WHO grade in the CGGA. D) The age distribution of CGGA oligodendrogliomas without *HOXD12* overexpression was not significantly non-Gaussian.

ure 7.3D, Supplemental Figure A.20). When patients with *HOXD12* overexpression were removed from the set of TCGA oligodendrogliomas, the age distribution of the remaining oligodendrogliomas was no longer significantly non-Gaussian (p=1.0, Shapiro-Wilk), indicating our hypothesized older, more aggressive subgroup of oligodendrogliomas may be enriched for tumors that overexpress *HOXD12* (Figure 7.3E).

These results were supported by data from the CGGA. Older (≥45 years old) CGGA oligodendroglioma patients were significantly enriched for *HOXD12* overexpression (p=0.03, Fisher's) (Figure 7.4A, Supplemental Figure A.21). Furthermore, *HOXD12* overexpression was prognostic in the CGGA ($p < 0.001$, log-rank) and had a similar hazard ratio (HR) as

observed in the TCGA (TCGA HR=9.3, CGGA HR=8.4) (Figure 7.4B). *HOXD12* overexpression was independently prognostic of tumor grade in the CGGA but lost its significance when patient age was accounted for, possibly because the sample size (N=44) was relatively small for a multivariate analysis (Figure 7.4C). As was the case in the TCGA, when CGGA oligodendroglioma patients with *HOXD12* overexpression were removed, the age distribution of the remaining oligodendrogliomas was no longer significantly non-Gaussian (p=1.0, Shapiro-Wilk) (Figure 7.4D). Similar trends were found in the GLASS dataset, but tumor selection bias along with the low count of patients likely prevented significance from being reached (Supplemental Figure A.22).

HOXD12 **Hypermethylation is Age-Associated and Predictive of Poor Survival Independent of** ***HOXD12*** **Overexpression**

Given that DNA methylation and gene transcription are related, we sought to evaluate *HOXD12* methylation levels. We found that every Illumina 450K array DNA methylation probe associated with *HOXD12* (N=15) was significantly positively correlated with *HOXD12* expression in the TCGA, suggesting that *HOXD12* gene body hypermethylation is associated with high transcriptional activity [167] (Figure 7.5A). Accordingly, we hypothesized that *HOXD12* hypermethylation would portend poor survival in oligodendroglioma. To test this hypothesis, we defined a *HOXD12* methylation signature and set a threshold above which we deemed patients *HOXD12* hypermethylated. Additionally, we classified every *HOXD12*-associated methylation probe according to its genomic position relative to *HOXD12*'s transcription start site (TSS) to give our results biological context (Figure 7.5B).

Probes included in our *HOXD12* methylation signature were chosen by testing the optimal survival-stratifying methylation beta value threshold for each probe (Figure 7.5C) for age and multivariate survival associations in the TCGA. Notably, the only three probes that passed all such tests were the three probes located within *HOXD12*'s gene body (exons 1 and 2), a CpG region whose hypermethylation has been associated with gene overexpression and

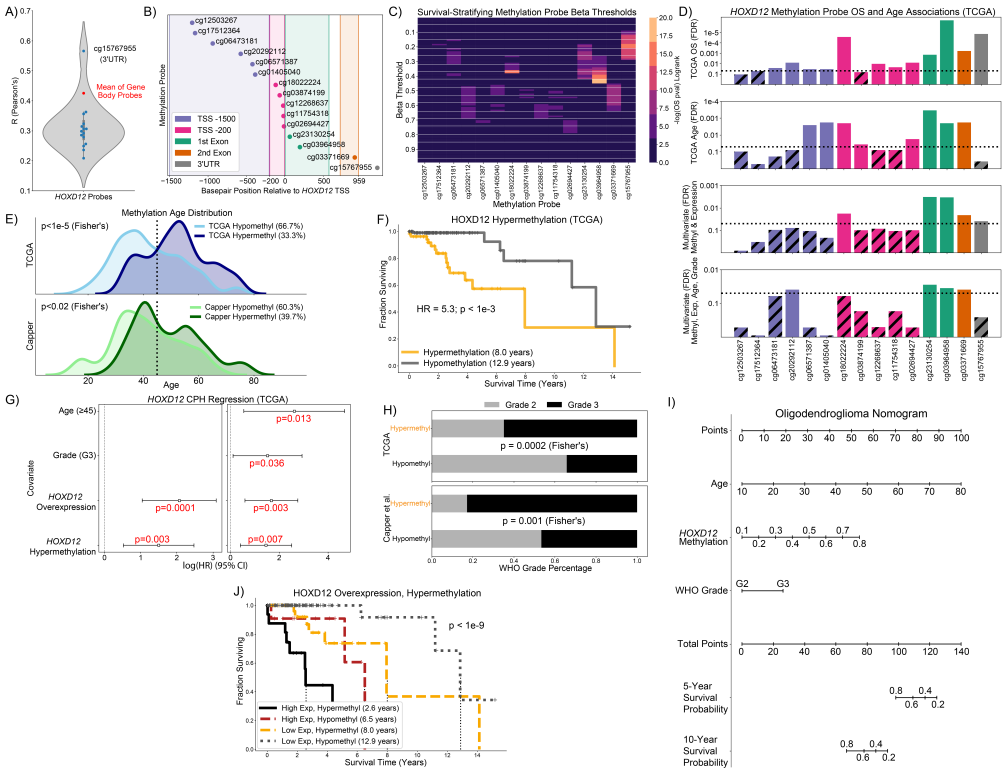


Figure 7.5: *HOXD12* Hypermethylation is Associated with Age and Predictive of Poor Survival Independent of *HOXD12* Overexpression. A) All *HOXD12*-associated methylation probes (N=15) were significantly positively correlated with *HOXD12* expression in the TCGA. B) *HOXD12*-associated methylation probes were classified according to their genomic position relative to *HOXD12*'s transcription start site (TSS). C) Beta value thresholds that best stratified survival in the TCGA for each *HOXD12*-associated probe were chosen for later inter-probe comparisons. D) *HOXD12*-associated probes were tested for age, univariate survival, and multivariate survival associations in the TCGA using specified thresholds. Only the three probes located within *HOXD12*'s gene body passed all such tests. E) *HOXD12* hypermethylation was associated with age in the TCGA and the Capper et al. cohort. *HOXD12* hypermethylation was prognostic in the TCGA in (F) univariate and (G) multivariate analyses. H) *HOXD12* hypermethylated oligodendrogliomas from Capper et al. and the TCGA were enriched for higher WHO grade. I) Nomogram for predicting 5- and 10-year overall survival using patient age, *HOXD12* methylation level, and WHO grade. J) Oligodendroglioma patients harboring overexpression and hypermethylation of *HOXD12* in the TCGA had the worst outcomes (overall survival = 2.6 years).

possibly tumorigenesis in other genes [168] (Figure 7.5D). We defined our *HOXD12* methylation signature as the mean of the three selected probes in *HOXD12*'s gene body and set the hypermethylation threshold as the threshold (0.358) that split the top third of TCGA oligodendroglioma patients from the bottom two thirds.

Using this definition for hypermethylation, we found that *HOXD12* hypermethylation was associated with age in the TCGA ($p<1e-5$, Fisher's) and Capper et al. cohorts ($p<0.02$, Fisher's) (Figure 7.5E). Furthermore, *HOXD12* hypermethylation was prognostic in the TCGA in a univariate analysis (HR=4.2, $p<0.001$, log-rank) and two multivariate analyses which accounted for *HOXD12* overexpression ($p<0.003$, CPH) and *HOXD12* overexpression, patient age, and tumor grade ($p<0.007$, CPH) (Figure 7.5F, G). Although outcome data was not available for the Capper et al. cohort, *HOXD12* hypermethylated. Oligodendrogliomas from Capper et al were substantially enriched for higher WHO grade (grade 3) ($p=0.001$, Fisher's), as were their TCGA counterparts ($p<0.001$, Fisher's), indicating that *HOXD12* hypermethylation was likely prognostic in the Capper et al. cohort (Figure 7.5H). Despite the small sample size in GLASS dataset, *HOXD12* hypermethylation was prognostic ($p<0.05$, log-rank) (Supplemental Figure A.23). For clinical applicability, we developed a nomogram predicting 5- and 10-year overall survival using patient age, *HOXD12* methylation level, and WHO grade using TCGA data (Figure 7.5I), and we tested whether immunohistochemistry for *HOXD12* could distinguish between *HOXD12*-hypermethylated and *HOXD12*-hypomethylated oligodendroglioma (Supplemental Figure A.24). Immunohistochemistry for *HOXD12* was not a sufficiently sensitive surrogate marker to detect any difference between *HOXD12* methylation levels. Overall, patients demonstrating overexpression and hypermethylation of *HOXD12* in the TCGA formed the class of oligodendrogliomas with the worst outcomes with a median overall survival of only 2.6 years (Figure 7.5J).

HOXD12 Overexpression and Hypermethylation are Prognostic Independent of Key Histopathologic, Genomic, and Radiographic Features

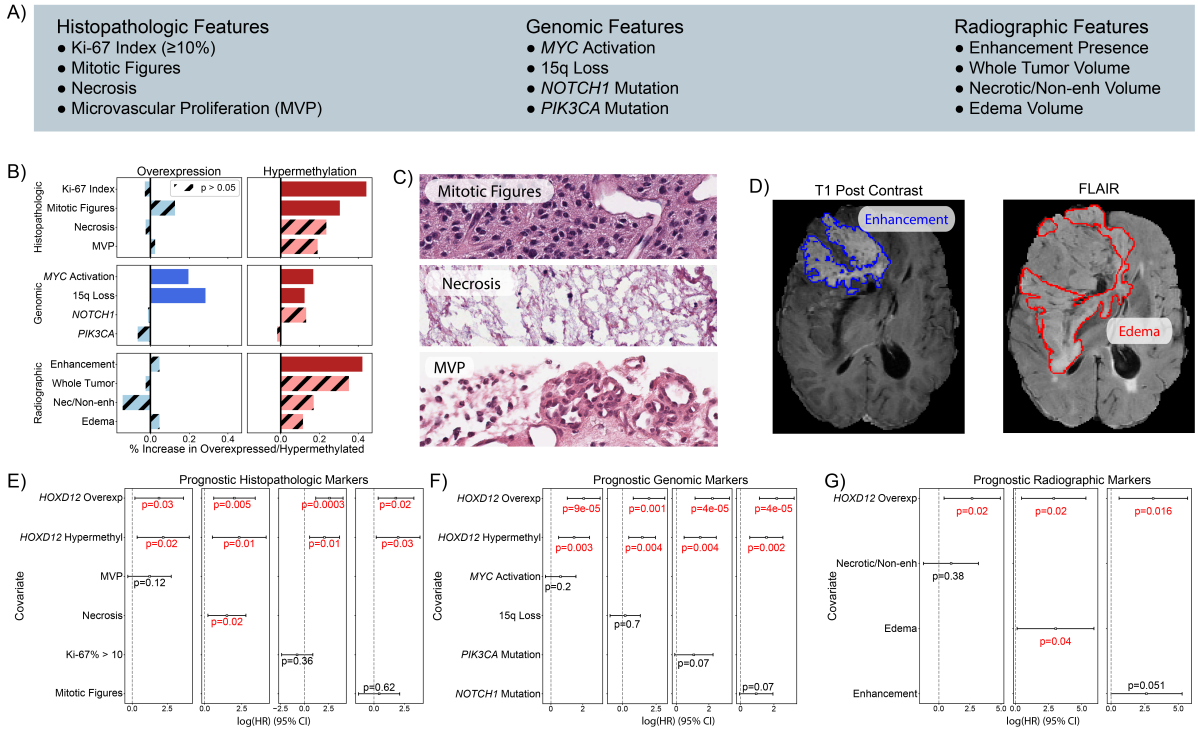


Figure 7.6: *HOXD12* Overexpression and Hypermethylation are Independently Prognostic of Key Histopathologic, Genomic, and Radiographic Features. A) List of key tested histopathologic, genomic, and radiographic features. B) *HOXD12* hypermethylation was linked to high Ki-67 proliferation index and the presence of mitotic figures. *HOXD12* overexpression and hypermethylation were associated with the loss of chromosome arm 15q and MYC activation. *HOXD12* hypermethylation was associated with the presence of T1-post contrast enhancement. C) Representative H&E-stained sections of a *HOXD12* overexpressing and hypermethylated TCGA oligodendroglioma (TCGA-DU-7018) showed increased mitotic activity, necrosis, and MVP. D) MRI images of the same patient show strong contrast enhancement (blue) and extensive peritumoral edema (red) on T1-post contrast and FLAIR MRI sequences, respectively. E,F) *HOXD12* overexpression and hypermethylation were independently prognostic of all tested histopathologic and genomic variables. G) *HOXD12* overexpression was independently prognostic of all tested radiographic features.

Having shown that *HOXD12* overexpression and hypermethylation were markers for poor survival in oligodendroglioma, we interrogated their relationship with clinically relevant histopathologic, genomic, and radiographic features in the TCGA (Figure 7.6A). Among histopathologic features, *HOXD12* hypermethylation was linked to high Ki-67 proliferation index ($\geq$10%) ($p<0.0001$, Fisher's) and the presence of mitotic figures ($p=0.015$, Fisher's), but not to the presence of necrosis or microvascular proliferation (MVP) (Figure 7.6B). Among proposed prognostic genomic biomarkers, we found that *HOXD12* overexpression and hypermethylation were associated with the loss of chromosome arm 15q ($p<0.001$, $p=0.049$, Fisher's) and *MYC* activation ($p=0.04$, $p<0.03$, Fisher's) but not with *NOTCH1* mutations or *PIK3CA* mutations. *CDKN2A* homozygous deletions were not tested, because fewer than 1% (1/171) of TCGA oligodendrogliomas harbor these alterations. Radiographically, *HOXD12* hypermethylation was associated with the presence of T1-post contrast enhancement ($p=0.03$, Fisher's), while *HOXD12* overexpression was not significantly associated with any tested radiographic feature. These associations suggest that tumors harboring *HOXD12* overexpression and hypermethylation are not only prognostic, but they also appear more aggressive, histopathologically, molecularly, and radiographically. For example, mitotic figures, necrosis, and MVP (Figure 7.6C), as well as contrast enhancement and substantial peritumoral edema (Figure 7.6D), were visible on a TCGA oligodendroglioma patient harboring both *HOXD12* overexpression and hypermethylation.

Remarkably, *HOXD12* overexpression and hypermethylation were independently prognostic of all tested histopathologic and genomic variables (Figure 7.6E, F). These multivariate analyses indicated that the statuses of *HOXD12* expression and methylation may be better indicators of outcome than histopathologic features used to determine WHO grade (mitotic features, necrosis, MVP) and genomic markers nominated by other studies as possible prognostic biomarkers [154, 155, 160, 56, 161]. Similarly, *HOXD12* overexpression was independently prognostic of all testable radiographic features (total tumor volume omitted due to collinearity); however, *HOXD12* hypermethylation was not, possibly because the

MRI dataset was small (N=27) (Figure 7.6G, Supplemental Figure A.25). Importantly, the survival associations observed were likely not due to treatment differences, because radiation and chemotherapy were evenly distributed between patients with or without *HOXD12* hypermethylation or overexpression (Supplemental Figure A.26). Regrettably, it was not possible to validate our findings in the CGGA or Capper et al. cohorts because of insufficient imaging or histopathology data.

TCGA Pan-cancer Analyses of *HOXD12* Overexpression and Hypermethylation Show Prognostic Value in Other Cancer Types

We sought to test whether the correlation, age association, and prognostic implications of *HOXD12* overexpression and hypermethylation we observed in oligodendroglioma held in other adult diffuse glioma subtypes and non-CNS cancer types. We found that *HOXD12* expression and methylation levels were significantly correlated in IDH-mutant astrocytomas and IDH-wildtype glioblastoma (Supplemental Figure A.27A), but that only in IDH-mutant astrocytomas were *HOXD12* overexpression and *HOXD12* hypermethylation associated with age ($p<0.05$, $p<0.001$, Fisher's) and survival (HR=3.0, $p<0.001$; HR=2.8, $p<0.001$, log-rank) (Supplemental Figure A.27B,C,D). In multivariate survival analyses, *HOXD12* hypermethylation ($p=0.02$, CPH), but not *HOXD12* overexpression ($p=0.2$, CPH), was independently prognostic of patient age in IDH-mutant astrocytoma, and neither were independently prognostic of patient age in IDH-wildtype glioblastoma.

In a pan-cancer analysis, we observed correlation between *HOXD12* expression and methylation levels (FDR < 0.05) in 24 of the 33 TCGA cancer types (73%) (Supplemental Figure A.28A). Age was not associated with both *HOXD12* overexpression and hypermethylation in any cancer type other than the TCGA lower-grade glioma (LGG) dataset, which contains virtually all TCGA oligodendrogliomas and most IDH-mutant astrocytomas (Supplemental Figure A.28B). *HOXD12* overexpression was prognostic in TCGA adrenocortical carcinoma (ACC FDR < 0.01, log-rank) and kidney renal papillary cell carcinoma (KIRP

FDR<1e-4, log-rank) in addition to low-grade gliomas (LGG FDR<1e-14, log-rank) (Supplemental Figure A.28C). Furthermore, *HOXD12* hypermethylation was also negatively prognostic in each of these cancer types (ACC FDR<1e-6, KIRP FDR<0.002, LGG FDR<1e-9, log-rank) along with kidney renal clear cell carcinoma (KIRC FDR<0.005), mesothelioma (MESO FDR<0.005), and sarcoma (SARC, FDR=0.029) (Supplemental Figure A.28D). Survival was driven primarily by *HOXD12* hypermethylation in ACC tumors, while *HOXD12* hypermethylation and overexpression were independently prognostic of each other and age in KIRP tumors (Supplemental Figure A.28E). In general, *HOXD12* expression and methylation levels were correlated and prognostic in selected cancer types, but not associated with patient age.

7.4 Discussion

Although patients diagnosed with oligodendroglioma typically have a better prognosis than their IDH-mutant astrocytoma and IDH-wildtype glioblastoma counterparts, the need for precision medicine in oligodendroglioma is important as it may influence treatment decisions. Oligodendrogliomas have more variable outcomes than other adult diffuse gliomas, which underscores the need to make therapy intensities proportionate to their risk. However, assessing a patient's risk is difficult as only a handful of oligodendroglioma risk stratifying biomarkers have been proposed to date [154, 155, 160, 56, 161]. Furthermore, the broadly held belief that oligodendroglioma is a single IDH-mutant and 1p/19q-codeleted entity has limited investigations of oligodendroglioma molecular heterogeneity and its possible prognostic implications [42, 46, 47, 45]. In this study, we addressed the need for novel biomarker discovery and challenged the notion that oligodendroglioma is a single entity. We propose that *HOXD12* overexpression and hypermethylation as poor prognostic biomarkers that may characterize two age and survival-associated subgroups of oligodendroglioma.

The few studies that have sought to interrogate oligodendroglioma molecular heterogeneity support our results. Kamoun et al. identified three subgroups of oligodendroglioma

from four microarray datasets, of which the largest subset was negatively prognostic in the TCGA (p=0.001, log-rank), weakly associated with age (p=0.08, t-test), and characterized by MYC activation [161]. Interestingly, we showed that TCGA patients that harbored either *HOXD12* overexpression or hypermethylation were associated with age and MYC activation, but *HOXD12* overexpression and hypermethylation were prognostic independent of both. This suggests that *HOXD12* overexpression and hypermethylation may explain the reported prognostic value of MYC activation and may be superior biomarkers. A single cell RNA-seq study by Tirosh et al. also interrogated oligodendroglioma molecular heterogeneity, albeit in an intra-tumoral setting, and linked neurodevelopmental transcription factor expression to a potentially proliferative stem/progenitor program in a small subpopulation of oligodendroglioma cells [169]. It may be that *HOXD12*, one of several developmental transcription factors upregulated in older oligodendrogliomas, is overexpressed in a larger proportion of cells that possess a stem/progenitor program that drives tumor growth. This hypothesis is supported by the absence of *HOXD12* expression in all 4,347 cells profiled by Tirosh et al. in six likely indolent WHO grade 2 oligodendroglioma tumors (data not shown).

This study intentionally focused on a single gene, but the explanation of atypically aggressive oligodendroglioma is likely far more complex than *HOXD12*'s molecular profile. Numerous genes in the HOX gene family have been reported to be involved in cancer (as reviewed by Gonçalves et al. [170]), particularly in glioblastoma [96, 170], primarily as oncogenes. However, in oligodendroglioma, HOX genes are unstudied beyond their capacity to differentiate oligodendroglial tumors from astrocytic tumors [171, 172]. Still, published evidence of HOX gene activity in cancer along with our observations that *HOXD12* was prognostic in several TCGA cancer types and that multiple HOX genes were upregulated in older oligodendroglioma patients, suggest that a collection of HOX genes may cooperate to drive an aggressive subtype of oligodendroglioma larger than what we reported in this study [169, 170]. Future work may broaden our analyses to all HOX genes and to other developmental transcription factors.

We were unable to functionally assess the biological role that *HOXD12* plays in oligo-

dendroglioma, the primary limitation of this work. Unfortunately, there are no suitable genetically engineered mouse models capable of supporting adequate oligodendroglioma animal models, and mouse xenografts, the best alternative, are exceedingly difficult to develop [160, 173, 174]. We also failed to find a consistent, convincing biological mechanism in the existing literature. *HOXD12* was reported in a nine-gene signature shown to predict survival in TCGA colon adenocarcinoma, but its role is unknown [175]. Other studies of *HOXD12* in glioma are limited to showing that *HOXD12* is upregulated in glioblastoma and pediatric low-grade gliomas compared to normal brain [176, 177]. Our results indicate that *HOXD12* is a promising prognostic biomarker, potentially superior to standard histopathologic and radiographic features as well as other proposed genomic biomarkers, but prospective human data must be collected and subjected to long-term follow up to confirm any findings and subsequent use in risk stratification and clinical decision-making.

Chapter 8

CONCLUSION

Adult-type diffuse gliomas are the most common primary adult malignant brain tumors and are universally regarded as lethal. Recent updates to the WHO classification of tumors of the CNS have used molecular biomarkers to define specific types of adult-type diffuse glioma and have led to more accurate diagnoses. However, patient outcomes still vary widely within each type. Novel prognostic biomarkers can improve the clinical management of patients and clinical trials via improved risk stratification. Survival-associated biomarker discovery is of utmost importance in adult-type diffuse glioma.

In chapter 3, we described an efficient neural network architecture, 3D-ESPNet, trained to predict adult-type diffuse glioma tumor segmentation masks that we used in subsequent biomarker discovery studies. We trained our segmentation model on the BraTS dataset, where it proved to be accurate in lower and higher-grade adult-type diffuse glioma. This segmentation network automated a crucial step in standard radiomic pipelines.

In chapter 4, we sought to establish a prognostic radiographic signature by predicting the status of prognostic genomic biomarkers in IDH-wildtype glioblastoma from MRI data. We devised a radiomic pipeline that preprocessed IDH-wildtype glioblastoma MRI data, segmented the tumors into four distinct compartments, and extracted radiomic features from each tumor compartment. We developed a novel LASSO-based feature selection method that boosted our prediction performance beyond traditional feature selection methods and hand-crafted features used by neuroradiologists. An examination of the most commonly selected features indicated that areas of peritumoral edema visible on T2-weighted MRI sequences were the most discriminative of prognostic SCNA profiles in IDH-wildtype glioblastoma.

In chapter 5, we used manifold learning and multitask learning (MTL) to leverage unla-

beled and partially labeled adult-type adult diffuse glioma MRI and SCNA data to improve survival and biomarker prediction. Tumor segmentation was used as an auxiliary learning task in MTL models learning IDH-mutation or 1p/19q-codeletion status. Jointly learning segmentation masks with biomarkers yielded better biomarker prediction and created MRI embeddings more capable of predicting patient survival.

In chapter 6, we assembled a large SCNA dataset and trained classification models to predict adult-type diffuse glioma molecular diagnoses from antiquated SCNA data collected before IDH mutation and 1p/19q-codeletion statuses were routinely collected. We achieved remarkably accurate validation results on modern datasets with gold-standard ground truth labels and observed the striking survival differences we expected between predicted adult-type diffuse glioma types in older retrospective datasets. Further, this study proposed rigorous thresholds for calling 1p/19q-codeletions and simultaneous whole gain of chromosome 7 and whole loss of chromosome 10 (+7/-10).

In chapter 7, we nominated *HOXD12* as a potential oncogene in oligodendroglioma that may drive a more aggressive subgroup of patients that occurs disproportionately in patients at an older age. *HOXD12* expression and *HOXD12* gene body hypermethylation were linked to older oligodendroglioma patients with shorter outcomes. In addition, *HOXD12* overexpression and gene body hypermethylation were shown to be independently prognostic of known survival-associated histologic, radiographic, and genomic biomarkers. This study has numerous clinical applications, including recommending milder treatment for those oligodendroglioma patients with low *HOXD12* expression and gene body methylation levels.

In the future, we will extend our work on *HOXD12* and apply the methods described in Chapter 7 to other cancer types. We will also investigate whether the observed *HOXD12* age and survival associations in oligodendroglioma hold in IDH-mutant tumors in general and assess other HOX genes for similar behaviors. We will also acquire functional data from oligodendroglioma human cell lines to test our hypotheses in vitro. Going forward, we will also link our genomic findings to more radiomic features that can be assessed over time via non-invasive MRI. Our work will continue to be focused on clinical translation.

BIBLIOGRAPHY

[1] M. L. Goodenberger and R. B. Jenkins, "Genetics of adult glioma," *Cancer genetics*, vol. 205, no. 12, 2012.

[2] P. Y. Wen, M. Weller, E. Q. Lee, B. M. Alexander, J. S. Barnholtz-Sloan, F. P. Barthel, T. T. Batchelor, R. S. Bindra, S. M. Chang, E. A. Chiocca, *et al.*, "Glioblastoma in adults: a society for neuro-oncology (sno) and european society of neuro-oncology (eano) consensus review on current management and future directions," *Neuro-oncology*, vol. 22, no. 8, pp. 1073–1113, 2020.

[3] P. Chuntova, F. Chow, P. B. Watchmaker, M. Galvez, A. B. Heimberger, E. W. Newell, A. Diaz, R. A. DePinho, M. O. Li, E. J. Wherry, *et al.*, "Unique challenges for glioblastoma immunotherapy—discussions across neuro-oncology and non-neuro-oncology experts in cancer immunology. meeting report from the 2019 sno immuno-oncology think tank," *Neuro-oncology*, vol. 23, no. 3, pp. 356–375, 2021.

[4] R. Medikonda, G. Dunn, M. Rahman, P. Fecci, and M. Lim, "A review of glioblastoma immunotherapy," *Journal of neuro-oncology*, vol. 151, no. 1, pp. 41–53, 2021.

[5] D. N. Louis, A. Perry, G. Reifenberger, A. Von Deimling, D. Figarella-Branger, W. K. Cavenee, H. Ohgaki, O. D. Wiestler, P. Kleihues, and D. W. Ellison, "The 2016 world health organization classification of tumors of the central nervous system: a summary," *Acta neuropathologica*, vol. 131, no. 6, pp. 803–820, 2016.

[6] D. N. Louis, A. Perry, P. Wesseling, D. J. Brat, I. A. Cree, D. Figarella-Branger, C. Hawkins, H. Ng, S. M. Pfister, G. Reifenberger, *et al.*, "The 2021 who classification of tumors of the central nervous system: a summary," *Neuro-oncology*, vol. 23, no. 8, pp. 1231–1251, 2021.

[7] N. Nuechterlein, L. G. Shapiro, E. C. Holland, and P. J. Cimino, "Machine learning modeling of genome-wide copy number alteration signatures reliably predicts idh mutational status in adult diffuse glioma," *Acta neuropathologica communications*, vol. 9, no. 1, pp. 1–18, 2021.

[8] R. Stupp, W. P. Mason, M. J. Van Den Bent, M. Weller, B. Fisher, M. J. Taphoorn, K. Belanger, A. A. Brandes, C. Marosi, U. Bogdahn, *et al.*, "Radiotherapy plus

concomitant and adjuvant temozolomide for glioblastoma," *New England journal of medicine*, vol. 352, no. 10, pp. 987–996, 2005.

[9] M. Weller, M. van den Bent, M. Preusser, E. Le Rhun, J. C. Tonn, G. Minniti, M. Bendszus, C. Balana, O. Chinot, L. Dirven, *et al.*, "Eano guidelines on the diagnosis and treatment of diffuse gliomas of adulthood," *Nature reviews Clinical oncology*, vol. 18, no. 3, pp. 170–186, 2021.

[10] J. G. Nicholson and H. A. Fine, "Diffuse glioma heterogeneity and its therapeutic implicationsdiffuse glioma heterogeneity and its therapeutic implications," *Cancer discovery*, vol. 11, no. 3, pp. 575–590, 2021.

[11] R. K. Oberoi, K. E. Parrish, T. T. Sio, R. K. Mittapalli, W. F. Elmquist, and J. N. Sarkaria, "Strategies to improve delivery of anticancer drugs across the blood–brain barrier to treat glioblastoma," *Neuro-oncology*, vol. 18, no. 1, pp. 27–36, 2015.

[12] D. J. Brat, K. Aldape, H. Colman, E. C. Holland, D. N. Louis, R. B. Jenkins, B. Kleinschmidt-DeMasters, A. Perry, G. Reifenberger, R. Stupp, *et al.*, "cimpact-now update 3: recommended diagnostic criteria for "diffuse astrocytic glioma, idh-wildtype, with molecular features of glioblastoma, who grade iv"," *Acta neuropathologica*, vol. 136, no. 5, pp. 805–810, 2018.

[13] R. A. Yoda, T. Marxen, L. Longo, C. Ene, H.-G. Wirsching, C. D. Keene, E. C. Holland, and P. J. Cimino, "Mitotic index thresholds do not predict clinical outcome for idh-mutant astrocytoma," *Journal of Neuropathology & Experimental Neurology*, vol. 78, no. 11, pp. 1002–1010, 2019.

[14] M. Shirahata, T. Ono, D. Stichel, D. Schrimpf, D. E. Reuss, F. Sahm, C. Koelsche, A. Wefers, A. Reinhardt, K. Huang, *et al.*, "Novel, improved grading system (s) for idh-mutant astrocytic gliomas," *Acta neuropathologica*, vol. 136, no. 1, pp. 153–166, 2018.

[15] P. J. Cimino, L. McFerrin, H.-G. Wirsching, S. Arora, H. Bolouri, R. Rabadan, M. Weller, and E. C. Holland, "Copy number profiling across glioblastoma populations has implications for clinical trial design," *Neuro-oncology*, vol. 20, no. 10, pp. 1368–1373, 2018.

[16] N. Nuechterlein, B. Li, A. Feroze, E. C. Holland, L. Shapiro, D. Haynor, J. Fink, and P. J. Cimino, "Radiogenomic modeling predicts survival-associated prognostic groups in glioblastoma," *Neuro-oncology advances*, vol. 3, no. 1, p. vdab004, 2021.

[17] S. J. Bagley, S. Kothari, R. Rahman, E. Q. Lee, G. P. Dunn, E. Galanis, S. M. Chang, L. B. Nabors, M. S. Ahluwalia, R. Stupp, *et al.*, "Glioblastoma clinical trials: Current landscape and opportunities for improvementcurrent glioblastoma clinical trial landscape," *Clinical Cancer Research*, vol. 28, no. 4, pp. 594–602, 2022.

[18] D. J. Brat, K. Aldape, H. Colman, D. Figrarella-Branger, G. N. Fuller, C. Giannini, E. C. Holland, R. B. Jenkins, B. Kleinschmidt-DeMasters, T. Komori, *et al.*, "cimpact-now update 5: recommended grading criteria and terminologies for idh-mutant astrocytomas," *Acta neuropathologica*, vol. 139, no. 3, pp. 603–608, 2020.

[19] D. N. Louis, P. Wesseling, K. Aldape, D. J. Brat, D. Capper, I. A. Cree, C. Eberhart, D. Figarella-Branger, M. Fouladi, G. N. Fuller, *et al.*, "cimpact-now update 6: new entity and diagnostic principle recommendations of the cimpact-utrecht meeting on future cns tumor classification and grading," 2020.

[20] D. A. Orringer, B. Pandian, Y. S. Niknafs, T. C. Hollon, J. Boyle, S. Lewis, M. Garrard, S. L. Hervey-Jumper, H. J. Garton, C. O. Maher, *et al.*, "Rapid intraoperative histology of unprocessed surgical specimens via fibre-laser-based stimulated raman scattering microscopy," *Nature biomedical engineering*, vol. 1, no. 2, pp. 1–13, 2017.

[21] T. C. Hollon, B. Pandian, A. R. Adapa, E. Urias, A. V. Save, S. S. S. Khalsa, D. G. Eichberg, R. S. D'Amico, Z. U. Farooq, S. Lewis, *et al.*, "Near real-time intraoperative brain tumor diagnosis using stimulated raman histology and deep neural networks," *Nature medicine*, vol. 26, no. 1, pp. 52–58, 2020.

[22] T. C. Hollon, B. Pandian, E. Urias, A. V. Save, A. R. Adapa, S. Srinivasan, N. K. Jairath, Z. Farooq, T. Marie, W. N. Al-Holou, *et al.*, "Rapid, label-free detection of diffuse glioma recurrence using intraoperative stimulated raman histology and deep neural networks," *Neuro-oncology*, vol. 23, no. 1, pp. 144–155, 2021.

[23] R. J. Gillies *et al.*, "Radiomics: images are more than pictures, they are data," *Radiology*, vol. 278, no. 2, 2016.

[24] V. Kumar *et al.*, "Radiomics: the process and the challenges," *Magnetic resonance imaging*, vol. 30, no. 9, 2012.

[25] Y. S. Choi, S. Bae, J. H. Chang, S.-G. Kang, S. H. Kim, J. Kim, T. H. Rim, S. H. Choi, R. Jain, and S.-K. Lee, "Fully automated hybrid approach to predict the idh mutation status of gliomas via deep learning and radiomics," *Neuro-oncology*, vol. 23, no. 2, pp. 304–313, 2021.

[26] Y. Matsui *et al.*, "Prediction of lower-grade glioma molecular subtypes using deep learning," *Journal of Neuro-Oncology*, vol. 146, no. 2, 2020.

[27] H. Arita *et al.*, "Lesion location implemented magnetic resonance imaging radiomics for predicting idh and tert promoter mutations in grade ii/iii gliomas," *Scientific reports*, vol. 8, no. 1, 2018.

[28] X. Zhang *et al.*, "Radiomics strategy for molecular subtype stratification of lower-grade glioma: detecting idh and tp53 mutations based on multimodal mri," *Journal of Magnetic Resonance Imaging*, vol. 48, no. 4, 2018.

[29] Z.-C. Li *et al.*, "Multiregional radiomics profiling from multiparametric mri: Identifying an imaging predictor of idh1 mutation status in glioblastoma," *Cancer medicine*, vol. 7, no. 12, 2018.

[30] P. Chang *et al.*, "Deep-learning convolutional neural networks accurately classify genetic mutations in gliomas," *American Journal of Neuroradiology*, vol. 39, no. 7, 2018.

[31] P. Korfiatis *et al.*, "Residual deep convolutional neural network predicts mgmt methylation status," *Journal of digital imaging*, vol. 30, no. 5, 2017.

[32] P. Kickingereder, D. Bonekamp, *et al.*, "Radiogenomics of glioblastoma: machine learning–based classification of molecular characteristics by using multiparametric and multiregional mr imaging features," *Radiology*, vol. 281, no. 3, pp. 907–918, 2016.

[33] P. Kickingereder, F. Isensee, I. Tursunova, J. Petersen, U. Neuberger, D. Bonekamp, G. Brugnara, M. Schell, T. Kessler, M. Foltyn, *et al.*, "Automated quantitative tumour response assessment of mri in neuro-oncology with artificial neural networks: a multicentre, retrospective study," *The Lancet Oncology*, vol. 20, no. 5, pp. 728–740, 2019.

[34] S. Bakas *et al.*, "Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the brats challenge," *arXiv preprint arXiv:1811.02629*, 2018.

[35] S. Mehta, N. Nuechterlein, E. Mercan, B. Li, S. Nofallah, W. Wu, X. Lu, A. Caspi, M. Rastegari, J. Elmore, *et al.*, "Applications of the espnet architecture in medical imaging," in *State of the Art in Neural Networks and their Applications*, pp. 117–131, Elsevier, 2021.

[36] C. Davatzikos, "Machine learning in neuroimaging: Progress and challenges," *Neuroimage*, vol. 197, p. 652, 2019.

[37] M. J. Sheller, B. Edwards, G. A. Reina, J. Martin, S. Pati, A. Kotrotsou, M. Milchenko, W. Xu, D. Marcus, R. R. Colen, *et al.*, "Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data," *Scientific reports*, vol. 10, no. 1, pp. 1–12, 2020.

[38] C. G. A. R. Network *et al.*, "Comprehensive genomic characterization defines human glioblastoma genes and core pathways," *Nature*, vol. 455, no. 7216, p. 1061, 2008.

[39] H. Noushmehr, D. J. Weisenberger, K. Diefes, H. S. Phillips, K. Pujara, B. P. Berman, F. Pan, C. E. Pelloski, E. P. Sulman, K. P. Bhat, *et al.*, "Identification of a cpg island methylator phenotype that defines a distinct subgroup of glioma," *Cancer cell*, vol. 17, no. 5, pp. 510–522, 2010.

[40] R. G. Verhaak, K. A. Hoadley, E. Purdom, V. Wang, Y. Qi, M. D. Wilkerson, C. R. Miller, L. Ding, T. Golub, J. P. Mesirov, *et al.*, "Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in pdgfra, idh1, egfr, and nf1," *Cancer cell*, vol. 17, no. 1, pp. 98–110, 2010.

[41] D. Sturm, H. Witt, V. Hovestadt, D.-A. Khuong-Quang, D. T. Jones, C. Konermann, E. Pfaff, M. Tönjes, M. Sill, S. Bender, *et al.*, "Hotspot mutations in h3f3a and idh1 define distinct epigenetic and biological subgroups of glioblastoma," *Cancer cell*, vol. 22, no. 4, pp. 425–437, 2012.

[42] C. G. A. R. Network, "Comprehensive, integrative genomic analysis of diffuse lower-grade gliomas," *New England Journal of Medicine*, vol. 372, no. 26, pp. 2481–2498, 2015.

[43] J. E. Eckel-Passow, D. H. Lachance, A. M. Molinaro, K. M. Walsh, P. A. Decker, H. Sicotte, M. Pekmezci, T. Rice, M. L. Kosel, I. V. Smirnov, *et al.*, "Glioma groups based on 1p/19q, idh, and tert promoter mutations in tumors," *New England Journal of Medicine*, vol. 372, no. 26, pp. 2499–2508, 2015.

[44] H. Suzuki, K. Aoki, K. Chiba, Y. Sato, Y. Shiozawa, Y. Shiraishi, T. Shimamura, A. Niida, K. Motomura, F. Ohka, *et al.*, "Mutational landscape and clonal architecture in grade ii and iii gliomas," *Nature genetics*, vol. 47, no. 5, pp. 458–468, 2015.

[45] M. Ceccarelli, F. P. Barthel, T. M. Malta, T. S. Sabedot, S. R. Salama, B. A. Murray, O. Morozova, Y. Newton, A. Radenbaugh, S. M. Pagnotta, *et al.*, "Molecular profiling reveals biologically discrete subsets and pathways of progression in diffuse glioma," *Cell*, vol. 164, no. 3, pp. 550–563, 2016.

[46] P. J. Cimino, M. Zager, L. McFerrin, H.-G. Wirsching, H. Bolouri, B. Hentschel, A. von Deimling, D. Jones, G. Reifenberger, M. Weller, *et al.*, "Multidimensional scaling of diffuse gliomas: application to the 2016 world health organization classification system with prognostically relevant molecular subtype discovery," *Acta neuropathologica communications*, vol. 5, no. 1, pp. 1–14, 2017.

[47] D. Capper, D. T. Jones, M. Sill, V. Hovestadt, D. Schrimpf, D. Sturm, C. Koelsche, F. Sahm, L. Chavez, D. E. Reuss, *et al.*, "Dna methylation-based classification of central nervous system tumours," *Nature*, vol. 555, no. 7697, pp. 469–474, 2018.

[48] L. Garofano, S. Migliozzi, Y. T. Oh, F. D'Angelo, R. D. Najac, A. Ko, B. Frangaj, F. P. Caruso, K. Yu, J. Yuan, *et al.*, "Pathway-based classification of glioblastoma uncovers a mitochondrial subtype with therapeutic vulnerabilities," *Nature cancer*, vol. 2, no. 2, pp. 141–156, 2021.

[49] P. J. Cimino, C. Ketchum, R. Turakulov, O. Singh, Z. Abdullaev, C. Giannini, P. Pytel, G. Y. Lopez, H. Colman, M. P. Nasrallah, *et al.*, "Expanded analysis of high-grade astrocytoma with piloid features identifies an epigenetically and clinically distinct subtype associated with neurofibromatosis type 1," *Acta neuropathologica*, pp. 1–12, 2022.

[50] M. Weller, R. G. Weber, E. Willscher, V. Riehmer, B. Hentschel, M. Kreuz, J. Felsberg, U. Beyer, H. Löffler-Wirth, K. Kaulich, *et al.*, "Molecular classification of diffuse cerebral who grade ii/iii gliomas using genome-and transcriptome-wide profiling improves stratification of prognostically distinct patient groups," *Acta neuropathologica*, vol. 129, no. 5, pp. 679–693, 2015.

[51] D. Stichel, A. Ebrahimi, D. Reuss, D. Schrimpf, T. Ono, M. Shirahata, G. Reifenberger, M. Weller, D. Hänggi, W. Wick, *et al.*, "Distribution of egfr amplification, combined chromosome 7 gain and chromosome 10 loss, and tert promoter mutation in brain tumors and their potential for the reclassification of idhwt astrocytoma to glioblastoma," *Acta neuropathologica*, vol. 136, no. 5, pp. 793–803, 2018.

[52] D. E. Reuss, A. Kratz, F. Sahm, D. Capper, D. Schrimpf, C. Koelsche, V. Hovestadt, M. Bewerunge-Hudler, D. T. Jones, J. Schittenhelm, *et al.*, "Adult idh wild type astrocytomas biologically and clinically resolve into other tumor entities," *Acta neuropathologica*, vol. 130, no. 3, pp. 407–417, 2015.

[53] C. W. Brennan, R. G. Verhaak, A. McKenna, B. Campos, H. Noushmehr, S. R. Salama, S. Zheng, D. Chakravarty, J. Z. Sanborn, S. H. Berman, *et al.*, "The somatic genomic landscape of glioblastoma," *Cell*, vol. 155, no. 2, pp. 462–477, 2013.

[54] A. P. Patel, I. Tirosh, J. J. Trombetta, A. K. Shalek, S. M. Gillespie, H. Wakimoto, D. P. Cahill, B. V. Nahed, W. T. Curry, R. L. Martuza, *et al.*, "Single-cell rna-seq highlights intratumoral heterogeneity in primary glioblastoma," *Science*, vol. 344, no. 6190, pp. 1396–1401, 2014.

[55] A. L. Rivera, C. E. Pelloski, M. R. Gilbert, H. Colman, C. De La Cruz, E. P. Sulman, B. N. Bekele, and K. D. Aldape, "Mgmt promoter methylation is predictive of response to radiotherapy and prognostic in the absence of adjuvant alkylating chemotherapy for glioblastoma," *Neuro-oncology*, vol. 12, no. 2, pp. 116–121, 2010.

[56] R. Appay, C. Dehais, C.-A. Maurage, A. Alentorn, C. Carpentier, C. Colin, F. Ducray, F. Escande, A. Idbaih, A. Kamoun, *et al.*, "Cdkn2a homozygous deletion is a strong adverse prognosis factor in diffuse malignant idh-mutant gliomas," *Neuro-oncology*, vol. 21, no. 12, pp. 1519–1528, 2019.

[57] N. Nuechterlein *et al.*, "3D-ESPNet with pyramidal refinement for volumetric brain tumor image segmentation," in *International MICCAI Brainlesion Workshop*, Springer, 2018.

[58] N. Nuechterlein, B. Li, M. S. Seyfioğlu, S. Mehta, P. J. Cimino, and L. Shapiro, "Leveraging unlabeled data for glioma molecular subtype and survival prediction," in *2020 25th International Conference on Pattern Recognition (ICPR)*, pp. 7149–7156, IEEE, 2021.

[59] N. Nuechterlein, S. Cimino, S. Arora, L. Shapiro, E. Holland, M. Gilbert, and P. Cimino, "Hoxd12 is an independently prognostic age-associated indicator for poor survival in oligodendroglioma," in *NEURO-ONCOLOGY*, vol. 24, pp. 15–15, OXFORD UNIV PRESS INC JOURNALS DEPT, 2001 EVANS RD, CARY, NC 27513 USA, 2022.

[60] B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, Y. Burren, N. Porz, J. Slotboom, R. Wiest, *et al.*, "The multimodal brain tumor image segmentation benchmark (brats)," *IEEE transactions on medical imaging*, vol. 34, no. 10, p. 1993, 2015.

[61] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. S. Kirby, J. B. Freymann, K. Farahani, and C. Davatzikos, "Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features," *Sci Data*, vol. 4, p. 170117, 09 2017.

[62] S. Bakas, M. Reyes, A. Jakab, S. Bauer, M. Rempfler, A. Crimi, R. T. Shinohara, C. Berger, S. M. Ha, M. Rozycki, B. Menze, *et al.*, "Identifying the best machine

learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the brats challenge," *arXiv preprint arXiv:1811.02629*, 2018.

[63] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. S. Kirby, J. B. Freymann, K. Farahani, and C. Davatzikos, "Segmentation labels and radiomic features for the pre-operative scans of the tcga-gbm collection," *The Cancer Imaging Archive*, 2017.

[64] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. S. Kirby, J. B. Freymann, K. Farahani, and C. Davatzikos, "Segmentation Labels and Radiomic Features for the Pre-operative Scans of the TCGA-LGG collection," *The Cancer Imaging Archive*, 2017.

[65] K. Clark, B. Vendt, K. Smith, J. Freymann, J. Kirby, P. Koppel, S. Moore, S. Phillips, D. Maffitt, M. Pringle, *et al.*, "The cancer imaging archive (tcia): maintaining and operating a public information repository," *Journal of digital imaging*, vol. 26, no. 6, pp. 1045–1057, 2013.

[66] M. Jenkinson, P. Bannister, M. Brady, and S. Smith, "Improved optimization for the robust and accurate linear registration and motion correction of brain images," *Neuroimage*, vol. 17, no. 2, pp. 825–841, 2002.

[67] M. Jenkinson and S. Smith, "A global optimisation method for robust affine registration of brain images," *Medical image analysis*, vol. 5, no. 2, pp. 143–156, 2001.

[68] C. H. Mermel, S. E. Schumacher, B. Hill, M. L. Meyerson, R. Beroukhim, and G. Getz, "Gistic2. 0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers," *Genome biology*, vol. 12, no. 4, pp. 1–14, 2011.

[69] M. J. Aryee, A. E. Jaffe, H. Corrada-Bravo, C. Ladd-Acosta, A. P. Feinberg, K. D. Hansen, and R. A. Irizarry, "Minfi: a flexible and comprehensive bioconductor package for the analysis of infinium dna methylation microarrays," *Bioinformatics*, vol. 30, no. 10, pp. 1363–1369, 2014.

[70] V. Hovestadt and M. Zapatka, "Conumee: enhanced copy-number variation analysis using illumina dna methylation arrays," *R package version*, vol. 1, no. 0, 2017.

[71] E. Cerami, J. Gao, U. Dogrusoz, B. E. Gross, S. O. Sumer, B. A. Aksoy, A. Jacobsen, C. J. Byrne, M. L. Heuer, E. Larsson, *et al.*, "The cbio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data," 2012.

[72] J. Gao, B. A. Aksoy, U. Dogrusoz, G. Dresdner, B. Gross, S. O. Sumer, Y. Sun, A. Jacobsen, R. Sinha, E. Larsson, *et al.*, "Integrative analysis of complex cancer genomics and clinical profiles using the cbioportal," *Science signaling*, vol. 6, no. 269, pp. pl1–pl1, 2013.

[73] P. Jonsson, A. L. Lin, R. J. Young, N. M. DiStefano, D. M. Hyman, B. T. Li, M. F. Berger, A. Zehir, M. Ladanyi, D. B. Solit, *et al.*, "Genomic correlates of disease progression and treatment response in prospectively characterized gliomas," *Clinical Cancer Research*, vol. 25, no. 18, pp. 5537–5547, 2019.

[74] F. P. Barthel, K. C. Johnson, F. S. Varn, A. D. Moskalik, G. Tanner, E. Kocakavuk, K. J. Anderson, O. Abiola, K. Aldape, K. D. Alfaro, *et al.*, "Longitudinal molecular trajectories of diffuse glioma in adults," *Nature*, vol. 576, no. 7785, pp. 112–120, 2019.

[75] "Glioma through the looking glass: molecular evolution of diffuse gliomas and the glioma longitudinal analysis consortium," *Neuro-oncology*, vol. 20, no. 7, pp. 873–884, 2018.

[76] Y. Gusev, K. Bhuvaneshwar, L. Song, J.-C. Zenklusen, H. Fine, and S. Madhavan, "The rembrandt study, a large collection of genomic data from brain cancer patients," *Scientific data*, vol. 5, no. 1, pp. 1–9, 2018.

[77] V. E. Seshan, A. B. Olshen, *et al.*, "Dnacopy: a package for analyzing dna copy data," *Bioconductor Vignette*, pp. 1–7, 2014.

[78] L. Collado-Torres, A. Nellore, K. Kammers, S. E. Ellis, M. A. Taub, K. D. Hansen, A. E. Jaffe, B. Langmead, and J. T. Leek, "Reproducible rna-seq analysis using recount2," *Nature biotechnology*, vol. 35, no. 4, pp. 319–321, 2017.

[79] S. Arora, S. S. Pattwell, E. C. Holland, and H. Bolouri, "Variability in estimated gene expression among commonly used rna-seq pipelines," *Scientific reports*, vol. 10, no. 1, pp. 1–9, 2020.

[80] Z. Zhao, K.-N. Zhang, Q. Wang, G. Li, F. Zeng, Y. Zhang, F. Wu, R. Chai, Z. Wang, C. Zhang, *et al.*, "Chinese glioma genome atlas (cgga): a comprehensive resource with functional genomic data from chinese glioma patients," *Genomics, proteomics & bioinformatics*, vol. 19, no. 1, pp. 1–12, 2021.

[81] L. Van der Maaten and G. Hinton, "Visualizing data using t-sne.," *Journal of machine learning research*, vol. 9, no. 11, 2008.

[82] M. J. Goldman, B. Craft, M. Hastie, K. Repečka, F. McDade, A. Kamath, A. Banerjee, Y. Luo, D. Rogers, A. N. Brooks, *et al.*, "Visualizing and interpreting cancer genomics data via the xena platform," *Nature biotechnology*, vol. 38, no. 6, pp. 675–678, 2020.

[83] K. Ellrott, M. H. Bailey, G. Saksena, K. R. Covington, C. Kandoth, C. Stewart, J. Hess, S. Ma, K. E. Chiotti, M. McLellan, *et al.*, "Scalable open science approach for mutation calling of tumor exomes using multiple genomic pipelines," *Cell systems*, vol. 6, no. 3, pp. 271–281, 2018.

[84] Y. Yuan, W. Xiang, L. Yanhui, L. Ruofei, L. Shuang, F. Yingjun, Z. Qiao, Y. Yanwu, and M. Qing, "Ki-67 overexpression in who grade ii gliomas is associated with poor postoperative seizure control," *Seizure*, vol. 22, no. 10, pp. 877–881, 2013.

[85] J. R. Fink, M. Muzi, M. Peck, and K. A. Krohn, "Multimodality Brain Tumor Imaging: MR Imaging, PET, and PET/MR Imaging," *J. Nucl. Med.*, vol. 56, pp. 1554–1561, Oct 2015.

[86] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, pp. 234–241, Springer, 2015.

[87] D. C. Ciresan, L. M. Gambardella, A. Giusti, and J. Schmidhuber, "Deep neural networks segment neuronal membranes in electron microscopy images," in *IN NIPS*, pp. 2852–2860, 2012.

[88] K. Kamnitsas, E. Ferrante, S. Parisot, C. Ledig, A. V. Nori, A. Criminisi, D. Rueckert, and B. Glocker, "Deepmedic for brain tumor segmentation," in *BrainLes@MICCAI*, 2016.

[89] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," *2016 Fourth International Conference on 3D Vision (3DV)*, pp. 565–571, 2016.

[90] S. Mehta *et al.*, "ESPNet: Efficient spatial pyramid of dilated convolutions for semantic segmentation," in *Proceedings of the european conference on computer vision (ECCV)*, 2018.

[91] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *IEEE transactions on pattern analysis and machine intelligence*, vol. 40, no. 4, pp. 834–848, 2018.

[92] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid scene parsing network," in *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 2881–2890, 2017.

[93] K. Kamnitsas, W. Bai, E. Ferrante, S. McDonagh, M. Sinclair, N. Pawlowski, M. Rajchl, M. Lee, B. Kainz, D. Rueckert, *et al.*, "Ensembles of multiple models and architectures for robust brain tumour segmentation," in *International MICCAI Brainlesion Workshop*, pp. 450–462, Springer, 2017.

[94] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[95] H. Bolouri, L. P. Zhao, and E. C. Holland, "Big data visualization identifies the multidimensional molecular landscape of human gliomas," *Proceedings of the national academy of sciences*, vol. 113, no. 19, pp. 5394–5399, 2016.

[96] P. J. Cimino, Y. Kim, H.-J. Wu, J. Alexander, H.-G. Wirsching, F. Szulzewsky, K. Pitter, T. Ozawa, J. Wang, J. Vazquez, *et al.*, "Increased hoxa5 expression provides a selective advantage for gain of whole chromosome 7 in idh wild-type glioblastoma," *Genes & development*, vol. 32, no. 7-8, pp. 512–523, 2018.

[97] C. H. Lee, B. O. Alpert, P. Sankaranarayanan, and O. Alter, "Gsvd comparison of patient-matched normal and tumor acgh profiles reveals global copy-number alterations predicting glioblastoma multiforme survival," *PloS one*, vol. 7, no. 1, p. e30098, 2012.

[98] S. Zheng, J. Fu, R. Vegesna, Y. Mao, L. E. Heathcock, W. Torres-Garcia, R. Ezhilarasan, S. Wang, A. McKenna, L. Chin, *et al.*, "A survey of intragenic breakpoints in glioblastoma identifies a distinct subset associated with poor survival," *Genes & development*, vol. 27, no. 13, pp. 1462–1472, 2013.

[99] C. Geisenberger, A. Mock, R. Warta, C. Rapp, C. Schwager, A. Korshunov, A.-K. Nied, D. Capper, B. Brors, C. Jungk, *et al.*, "Molecular profiling of long-term survivors identifies a subgroup of glioblastoma characterized by chromosome 19/20 co-gain," *Acta neuropathologica*, vol. 130, no. 3, pp. 419–434, 2015.

[100] J. J. Van Griethuysen, A. Fedorov, C. Parmar, A. Hosny, N. Aucoin, V. Narayan, R. G. Beets-Tan, J.-C. Fillion-Robin, S. Pieper, and H. J. Aerts, "Computational radiomics system to decode the radiographic phenotype," *Cancer research*, vol. 77, no. 21, pp. e104–e107, 2017.

[101] C. Davatzikos, S. Rathore, S. Bakas, S. Pati, M. Bergman, R. Kalarot, P. Sridharan, A. Gastounioti, N. Jahani, E. Cohen, *et al.*, "Cancer imaging phenomics toolkit: quantitative imaging analytics for precision diagnostics and predictive modeling of clinical outcome," *Journal of medical imaging*, vol. 5, no. 1, p. 011018, 2018.

[102] S. Bakas, H. Akbari, A. Sotiras, *et al.*, "Advancing the cancer genome atlas glioma mri collections with expert segmentation labels and radiomic features," *Scientific data*, vol. 4, 2017.

[103] B. H. Menze *et al.*, "The multimodal brain tumor image segmentation benchmark (brats)," *IEEE transactions on medical imaging*, vol. 34, no. 10, 2014.

[104] P. A. Yushkevich, J. Piven, H. C. Hazlett, R. G. Smith, S. Ho, J. C. Gee, and G. Gerig, "User-guided 3d active contour segmentation of anatomical structures: significantly improved efficiency and reliability," *Neuroimage*, vol. 31, no. 3, pp. 1116–1128, 2006.

[105] N. J. Tustison, B. B. Avants, P. A. Cook, Y. Zheng, A. Egan, P. A. Yushkevich, and J. C. Gee, "N4itk: improved n3 bias correction," *IEEE transactions on medical imaging*, vol. 29, no. 6, pp. 1310–1320, 2010.

[106] R. M. Haralick, "Statistical and structural approaches to texture," *Proceedings of the IEEE*, vol. 67, no. 5, pp. 786–804, 1979.

[107] A. Hair, "On the theory of orthogonal function systems.," *MathematicalAnnals*, vol. 69, no. 3, pp. 331–371, 1910.

[108] P. J. Burt and E. H. Adelson, "The laplacian pyramid as a compact image code," in *Readings in computer vision*, pp. 671–679, Elsevier, 1987.

[109] D. A. Gutman, L. A. Cooper, S. N. Hwang, C. A. Holder, J. Gao, T. D. Aurora, W. D. Dunn Jr, L. Scarpace, T. Mikkelsen, R. Jain, *et al.*, "Mr imaging predictors of molecular profile and survival: multi-institutional study of the tcga glioblastoma data set," *Radiology*, vol. 267, no. 2, pp. 560–569, 2013.

[110] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, *et al.*, "Scikit-learn: Machine learning in python. the journal of machine learning research 12," 2011.

[111] T. Chen, T. He, M. Benesty, V. Khotilovich, Y. Tang, H. Cho, *et al.*, "Xgboost: extreme gradient boosting," *R package version 0.4-2*, vol. 1, no. 4, 2015.

[112] S. Rathore, S. Bakas, *et al.*, "Brain cancer imaging phenomics toolkit (brain-captk): an interactive platform for quantitative analysis of glioblastoma," in *International MICCAI Brainlesion Workshop*, Springer, 2017.

[113] J. E. Park *et al.*, "A systematic review reporting quality of radiomics research in neuro-oncology: toward clinical utility and quality improvement using high-dimensional imaging features," *BMC cancer*, vol. 20, no. 1, 2020.

[114] K. Kamnitsas *et al.*, "Deepmedic for brain tumor segmentation," in *International workshop on Brainlesion: Glioma, multiple sclerosis, stroke and traumatic brain injuries*, Springer, 2016.

[115] S. Bakas *et al.*, "Segmentation labels and radiomic features for the pre-operative scans of the tcga-lgg collection," *The Cancer Imaging Archive*, vol. 286, 2017.

[116] S. Bakas, H. Akbari, *et al.*, "Segmentation labels and radiomic features for the pre-operative scans of the tcga-gbm collection. the cancer imaging archive," *Nat Sci Data*, vol. 4, p. 170117, 2017.

[117] R. Caruana, "Multitask learning," *Machine learning*, vol. 28, no. 1, pp. 41–75, 1997.

[118] T. Evgeniou *et al.*, "Regularized multi–task learning," in *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, 2004.

[119] K. He *et al.*, "Mask r-cnn," in *Proceedings of the IEEE international conference on computer vision*, 2017.

[120] E. Hajiramezanali *et al.*, "Bayesian multi-domain learning for cancer subtype discovery from next-generation sequencing count data," in *Advances in Neural Information Processing Systems*, pp. 9115–9124, 2018.

[121] O. Ronneberger *et al.*, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, Springer, 2015.

[122] Z. Li *et al.*, "Deep learning based radiomics (dlr) and its usage in noninvasive idh1 prediction for low grade glioma," *Scientific reports*, vol. 7, no. 1, 2017.

[123] C. Su *et al.*, "Radiomics based on multicontrast mri can precisely differentiate among glioma subtypes and predict tumour-proliferative behaviour," *European radiology*, vol. 29, no. 4, 2019.

[124] J. Lao *et al.*, "A deep learning-based radiomics model for prediction of survival in glioblastoma multiforme," *Scientific reports*, vol. 7, no. 1, 2017.

[125] M. Sundararajan, A. Taly, and Q. Yan, "Axiomatic attribution for deep networks," in *International Conference on Machine Learning*, pp. 3319–3328, PMLR, 2017.

[126] Q. T. Ostrom, N. Patil, G. Cioffi, K. Waite, C. Kruchko, and J. S. Barnholtz-Sloan, "Cbtrus statistical report: primary brain and other central nervous system tumors diagnosed in the united states in 2013–2017," *Neuro-oncology*, vol. 22, no. Supplement_1, pp. iv1–iv96, 2020.

[127] P. J. Cimino and E. C. Holland, "Targeted copy number analysis outperforms histologic grading in predicting patient survival for who grades ii/iii idh-mutant astrocytomas," *Neuro-oncology*, vol. 21, no. 6, pp. 819–821, 2019.

[128] D. N. Louis, H. Ohgaki, O. D. Wiestler, W. K. Cavenee, P. C. Burger, A. Jouvet, B. W. Scheithauer, and P. Kleihues, "The 2007 who classification of tumours of the central nervous system," *Acta neuropathologica*, vol. 114, no. 2, pp. 97–109, 2007.

[129] D. E. Reuss, Y. Mamatjan, D. Schrimpf, D. Capper, V. Hovestadt, A. Kratz, F. Sahm, C. Koelsche, A. Korshunov, A. Olar, *et al.*, "Idh mutant diffuse and anaplastic astrocytomas have similar age at presentation and little difference in survival: a grading problem for who," *Acta neuropathologica*, vol. 129, no. 6, pp. 867–873, 2015.

[130] H. Yan, D. W. Parsons, G. Jin, R. McLendon, B. A. Rasheed, W. Yuan, I. Kos, I. Batinic-Haberle, S. Jones, G. J. Riggins, *et al.*, "Idh1 and idh2 mutations in gliomas," *New England journal of medicine*, vol. 360, no. 8, pp. 765–773, 2009.

[131] J. H. Carter, S. N. McNulty, P. J. Cimino, C. E. Cottrell, J. W. Heusel, K. A. Vigh-Conrad, and E. J. Duncavage, "Targeted next-generation sequencing in molecular subtyping of lower-grade diffuse gliomas: application of the world health organization's 2016 revised criteria for central nervous system tumors," *The Journal of Molecular Diagnostics*, vol. 19, no. 2, pp. 328–337, 2017.

[132] J. Felsberg, M. Wolter, H. Seul, B. Friedensdorf, M. Göppert, M. C. Sabel, and G. Reifenberger, "Rapid and sensitive assessment of the idh1 and idh2 mutation status in cerebral gliomas based on dna pyrosequencing," *Acta neuropathologica*, vol. 119, no. 4, pp. 501–507, 2010.

[133] P. Setty, J. Hammes, T. Rothämel, V. Vladimirova, C. M. Kramm, T. Pietsch, and A. Waha, "A pyrosequencing-based assay for the rapid detection of idh1 mutations in clinical samples," *The Journal of Molecular Diagnostics*, vol. 12, no. 6, pp. 750–756, 2010.

[134] M. S. Waitkus, B. H. Diplas, and H. Yan, "Isocitrate dehydrogenase mutations in gliomas," *Neuro-oncology*, vol. 18, no. 1, pp. 16–26, 2015.

[135] A. Zacher, K. Kaulich, S. Stepanow, M. Wolter, K. Köhrer, J. Felsberg, B. Malzkorn, and G. Reifenberger, "Molecular diagnostics of gliomas using next generation sequencing of a glioma-tailored gene panel," *Brain Pathology*, vol. 27, no. 2, pp. 146–159, 2017.

[136] A. Lasocki, F. Gaillard, A. Gorelik, and M. Gonzales, "Mri features can predict 1p/19q status in intracranial gliomas," *American Journal of Neuroradiology*, vol. 39, no. 4, pp. 687–692, 2018.

[137] N. Nuechterlein, B. Li, L. Shapiro, P. Cimino, and E. Holland, "Genome-wide somatic copy number alteration signatures have prognostic implications for adult astrocytic glioma," in *NEURO-ONCOLOGY*, vol. 24, pp. 9–10, OXFORD UNIV PRESS INC JOURNALS DEPT, 2001 EVANS RD, CARY, NC 27513 USA, 2022.

[138] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, and R. Durbin, "The sequence alignment/map format and samtools," *Bioinformatics*, vol. 25, no. 16, pp. 2078–2079, 2009.

[139] A. R. Quinlan and I. M. Hall, "Bedtools: a flexible suite of utilities for comparing genomic features," *Bioinformatics*, vol. 26, no. 6, pp. 841–842, 2010.

[140] B. Langmead and S. L. Salzberg, "Fast gapped-read alignment with bowtie 2," *Nature methods*, vol. 9, no. 4, pp. 357–359, 2012.

[141] A. McKenna, M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, *et al.*, "The genome analysis toolkit: a mapreduce framework for analyzing next-generation dna sequencing data," *Genome research*, vol. 20, no. 9, pp. 1297–1303, 2010.

[142] V. A. Adalsteinsson, G. Ha, S. S. Freeman, A. D. Choudhury, D. G. Stover, H. A. Parsons, G. Gydush, S. C. Reed, D. Rotem, J. Rhoades, *et al.*, "Scalable whole-exome sequencing of cell-free dna reveals high concordance with metastatic tumors," *Nature communications*, vol. 8, no. 1, pp. 1–13, 2017.

[143] M. Ali, "Pycaret: An open source, low-code machine learning library in python," *PyCaret version*, vol. 2, 2020.

[144] L. McInnes, J. Healy, and J. Melville, "Umap: Uniform manifold approximation and projection for dimension reduction," *arXiv preprint arXiv:1802.03426*, 2018.

[145] S. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," *arXiv preprint arXiv:1705.07874*, 2017.

[146] D. Chicco and G. Jurman, "The advantages of the matthews correlation coefficient (mcc) over f1 score and accuracy in binary classification evaluation," *BMC genomics*, vol. 21, no. 1, pp. 1–13, 2020.

[147] R. B. Jenkins, H. Blair, K. V. Ballman, C. Giannini, R. M. Arusell, M. Law, H. Flynn, S. Passe, S. Felten, P. D. Brown, *et al.*, "A t (1; 19)(q10; p10) mediates the combined deletions of 1p and 19q and predicts a better prognosis of patients with oligodendroglioma," *Cancer research*, vol. 66, no. 20, pp. 9852–9861, 2006.

[148] C. A. Griffin, P. Burger, L. Morsberger, R. Yonescu, S. Swierczynski, J. D. Weingart, and K. M. Murphy, "Identification of der (1; 19)(q10; p10) in five oligodendrogliomas suggests mechanism of concurrent 1p and 19q loss," *Journal of Neuropathology & Experimental Neurology*, vol. 65, no. 10, pp. 988–994, 2006.

[149] A. Korshunov, D. Capper, D. Reuss, D. Schrimpf, M. Ryzhova, V. Hovestadt, D. Sturm, J. Meyer, C. Jones, O. Zheludkova, *et al.*, "Histologically distinct neuroepithelial tumors with histone 3 g34 mutation are molecularly similar and comprise a single nosologic entity," *Acta neuropathologica*, vol. 131, no. 1, pp. 137–146, 2016.

[150] C. E. Fuller, R. E. Schmidt, K. A. Roth, P. C. Burger, B. W. Scheithauer, R. Banerjee, K. Trinkaus, R. Lytle, and A. Perry, "Clinical utility of fluorescence in situ hybridization (fish) in morphologically ambiguous gliomas with hybrid oligodendroglial/astrocytic features," *Journal of Neuropathology & Experimental Neurology*, vol. 62, no. 11, pp. 1118–1128, 2003.

[151] F. Sahm, D. Reuss, C. Koelsche, D. Capper, J. Schittenhelm, S. Heim, D. T. Jones, S. M. Pfister, C. Herold-Mende, W. Wick, *et al.*, "Farewell to oligoastrocytoma: in situ molecular genetics favor classification as either oligodendroglioma or astrocytoma," *Acta neuropathologica*, vol. 128, no. 4, pp. 551–559, 2014.

[152] K. Galbraith, A. Kumar, K. G. Abdullah, J. M. Walker, S. H. Adams, T. Prior, R. Dimentberg, F. C. Henderson, K. Mirchia, A. A. Sathe, *et al.*, "Molecular correlates of long survival in idh-wildtype glioblastoma cohorts," *Journal of Neuropathology & Experimental Neurology*, vol. 79, no. 8, pp. 843–854, 2020.

[153] H. Arita, Y. Matsushita, R. Machida, K. Yamasaki, N. Hata, M. Ohno, S. Yamaguchi, T. Sasayama, S. Tanaka, F. Higuchi, *et al.*, "Tert promoter mutation confers favorable prognosis regardless of 1p/19q status in adult diffuse gliomas with idh1/2 mutations," *Acta neuropathologica communications*, vol. 8, no. 1, pp. 1–11, 2020.

[154] K. Aoki, H. Nakamura, H. Suzuki, K. Matsuo, K. Kataoka, T. Shimamura, K. Motomura, F. Ohka, S. Shiina, T. Yamamoto, *et al.*, "Prognostic relevance of genetic alterations in diffuse lower-grade gliomas," *Neuro-oncology*, vol. 20, no. 1, pp. 66–77, 2018.

[155] S. H. Halani, S. Yousefi, J. V. Vega, M. R. Rossi, Z. Zhao, F. Amrollahi, C. A. Holder, A. Baxter-Stoltzfus, J. Eschbacher, B. Griffith, *et al.*, "Multi-faceted computational assessment of risk and progression in oligodendroglioma implicates notch and pi3k pathways," *NPJ precision oncology*, vol. 2, no. 1, pp. 1–9, 2018.

[156] G. Zadeh, O. H. Khan, M. Vogelbaum, and D. Schiff, "Much debated controversies of diffuse low-grade gliomas," *Neuro-oncology*, vol. 17, no. 3, p. 323, 2015.

[157] K. A. Jaeckle, K. V. Ballman, M. Van Den Bent, C. Giannini, E. Galanis, P. D. Brown, R. B. Jenkins, J. G. Cairncross, W. Wick, M. Weller, *et al.*, "Codel: phase iii study of rt, rt+ tmz, or tmz for newly diagnosed 1p/19q codeleted oligodendroglioma. analysis from the initial study design," *Neuro-oncology*, vol. 23, no. 3, pp. 457–467, 2021.

[158] M. J. van den Bent and S. M. Chang, "Grade ii and iii oligodendroglioma and astrocytoma," *Neurologic clinics*, vol. 36, no. 3, pp. 467–484, 2018.

[159] J. Weller, S. Katzendobler, P. Karschnia, S. Lietke, R. Egensperger, N. Thon, M. Weller, B. Suchorska, and J.-C. Tonn, "Pcv chemotherapy alone for who grade 2 oligodendroglioma: prolonged disease control with low risk of malignant progression," *Journal of neuro-oncology*, vol. 153, no. 2, pp. 283–291, 2021.

[160] K. Tateishi, T. Nakamura, T. A. Juratli, E. A. Williams, Y. Matsushita, S. Miyake, M. Nishi, J. J. Miller, S. S. Tummala, A. L. Fink, *et al.*, "Pi3k/akt/mtor pathway alterations promote malignant progression and xenograft formation in oligodendroglial tumorspi3k pathway activation in progression of oligodendroglioma," *Clinical Cancer Research*, vol. 25, no. 14, pp. 4375–4387, 2019.

[161] A. Kamoun, A. Idbaih, C. Dehais, N. Elarouci, C. Carpentier, E. Letouzé, C. Colin, K. Mokhtari, A. Jouvet, E. Uro-Coste, *et al.*, "Integrated multi-omics analysis of oligodendroglial tumours identifies three subgroups of 1p/19q co-deleted gliomas," *Nature communications*, vol. 7, no. 1, pp. 1–11, 2016.

[162] A. L. Garton, C. J. Kinslow, A. I. Rae, A. Mehta, S. C. Pannullo, R. S. Magge, R. Ramakrishna, G. M. McKhann, M. B. Sisti, J. N. Bruce, *et al.*, "Extent of resection, molecular signature, and survival in 1p19q-codeleted gliomas," *Journal of neurosurgery*, vol. 134, no. 5, pp. 1357–1367, 2020.

[163] T. Ozawa, M. Riester, Y.-K. Cheng, J. T. Huse, M. Squatrito, K. Helmy, N. Charles, F. Michor, and E. C. Holland, "Most human non-gcimp glioblastoma subtypes evolve from a common proneural-like precursor glioma," *Cancer cell*, vol. 26, no. 2, pp. 288–300, 2014.

[164] M. I. Love, W. Huber, and S. Anders, "Moderated estimation of fold change and dispersion for rna-seq data with deseq2," *Genome biology*, vol. 15, no. 12, pp. 1–21, 2014.

[165] M. Pojo, C. S. Gonçalves, A. Xavier-Magalhães, A. I. Oliveira, T. Gonçalves, S. Correia, A. J. Rodrigues, S. Costa, L. Pinto, A. A. Pinto, *et al.*, "A transcriptomic signature mediated by hoxa9 promotes human glioblastoma initiation, aggressiveness and resistance to temozolomide," *Oncotarget*, vol. 6, no. 10, p. 7657, 2015.

[166] M. Bibikova, B. Barnes, C. Tsan, V. Ho, B. Klotzle, J. M. Le, D. Delano, L. Zhang, G. P. Schroth, K. L. Gunderson, *et al.*, "High density dna methylation array with single cpg site resolution," *Genomics*, vol. 98, no. 4, pp. 288–295, 2011.

[167] J. Smith, S. Sen, R. J. Weeks, M. R. Eccles, and A. Chatterjee, "Promoter dna hypermethylation and paradoxical gene activation," *Trends in cancer*, vol. 6, no. 5, pp. 392–406, 2020.

[168] X. Yang, H. Han, D. D. De Carvalho, F. D. Lay, P. A. Jones, and G. Liang, "Gene body methylation can alter gene expression and is a therapeutic target in cancer," *Cancer cell*, vol. 26, no. 4, pp. 577–590, 2014.

[169] I. Tirosh, A. S. Venteicher, C. Hebert, L. E. Escalante, A. P. Patel, K. Yizhak, J. M. Fisher, C. Rodman, C. Mount, M. G. Filbin, *et al.*, "Single-cell rna-seq supports a developmental hierarchy in human oligodendroglioma," *Nature*, vol. 539, no. 7628, pp. 309–313, 2016.

[170] C. S. Gonçalves, E. Le Boiteux, P. Arnaud, and B. M. Costa, "Hox gene cluster (de) regulation in brain: From neurodevelopment to malignant glial tumours," *Cellular and Molecular Life Sciences*, vol. 77, no. 19, pp. 3797–3821, 2020.

[171] C. Lauber, B. Klink, and M. Seifert, "Comparative analysis of histologically classified oligodendrogliomas reveals characteristic molecular differences between subgroups," *BMC cancer*, vol. 18, no. 1, pp. 1–16, 2018.

[172] X. Zhou, T. Liang, J. Deng, K. Ng, M. Li, C. Lv, J. Chen, K. Yang, Z. Ma, W. Ma, *et al.*, "Differential and prognostic significance of hoxb7 in gliomas," *Frontiers in Cell and Developmental Biology*, p. 2222, 2021.

[173] B. Klink, H. Miletic, D. Stieber, P. C. Huszthy, J. A. C. Valenzuela, J. Balss, J. Wang, M. Schubert, P. Ø. Sakariassen, T. Sundstrøm, *et al.*, "A novel, diffusely infiltrative xenograft model of human anaplastic oligodendroglioma with mutations in fubp1, cic, and idh1," *PloS one*, vol. 8, no. 3, p. e59773, 2013.

[174] N. D. Exner, J. A. C. Valenzuela, K. Abou-El-Ardat, H. Miletic, P. C. Huszthy, P. M. Radehaus, E. Schröck, R. Bjerkvig, L. Kaderali, B. Klink, *et al.*, "Deep sequencing of a recurrent oligodendroglioma and the derived xenografts reveals new insights into the evolution of human oligodendroglioma and candidate driver genes," *Oncotarget*, vol. 10, no. 38, p. 3641, 2019.

[175] Y. Lu, S. Wu, C. Cui, M. Yu, S. Wang, Y. Yue, M. Liu, and Z. Sun, "Gene expression along with genomic copy number variation and mutational analysis were used to develop a 9-gene signature for estimating prognosis of coad," *OncoTargets and therapy*, vol. 13, p. 10393, 2020.

[176] M. Gallo, J. Ho, F. J. Coutinho, R. Vanner, L. Lee, R. Head, E. K. Ling, I. D. Clarke, and P. B. Dirks, "A tumorigenic mll-homeobox network in human glioblastoma stem cellsan mll-homeobox network in glioblastoma," *Cancer research*, vol. 73, no. 1, pp. 417–427, 2013.

[177] A. M. Buccoliero, F. Castiglione, D. R. Degl'Innocenti, F. Ammanati, F. Giordano, M. Sanzo, F. Mussa, L. Genitori, and G. L. Taddei, "Hox-d genes expression in pediatric low-grade gliomas: real-time-pcr study," *Cellular and molecular neurobiology*, vol. 29, no. 1, pp. 1–6, 2009.

[178] B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, Y. Burren, N. Porz, J. Slotboom, R. Wiest, L. Lanczi, E. Gerstner, M. Weber, T. Arbel, B. B. Avants, N. Ayache, P. Buendia, D. L. Collins, N. Cordier, J. J. Corso, A. Criminisi, T. Das, H. Delingette, Demiralp, C. R. Durst, M. Dojat, S. Doyle, J. Festa, F. Forbes, E. Geremia, B. Glocker, P. Golland, X. Guo, A. Hamamci, K. M. Iftekharuddin, R. Jena, N. M. John, E. Konukoglu, D. Lashkari, J. A. Mariz, R. Meier, S. Pereira, D. Precup, S. J. Price, T. R. Raviv, S. M. S. Reza, M. Ryan, D. Sarikaya, L. Schwartz, H. Shin, J. Shotton, C. A. Silva, N. Sousa, N. K. Subbanna, G. Szekely, T. J. Taylor, O. M. Thomas, N. J. Tustison, G. Unal, F. Vasseur, M. Wintermark, D. H. Ye, L. Zhao, B. Zhao, D. Zikic, M. Prastawa, M. Reyes, and K. V. Leemput, "The multimodal brain tumor image segmentation benchmark (brats)," *IEEE Transactions on Medical Imaging*, vol. 34, pp. 1993–2024, Oct 2015.

[179] N. J. Tustison, B. B. Avants, P. A. Cook, Y. Zheng, A. Egan, P. A. Yushkevich, and J. C. Gee, "N4ITK: improved N3 bias correction," *IEEE Trans Med Imaging*, vol. 29, pp. 1310–1320, Jun 2010.

[180] S. Mehta, M. Rastegari, A. Caspi, L. Shapiro, and H. Hajishirzi, "Espnet: Efficient spatial pyramid of dilated convolutions for semantic segmentation," *arXiv preprint arXiv:1803.06815*, 2018.

[181] F. Isensee, P. Kickingereder, D. Bonekamp, M. Bendszus, W. Wick, H. Schlemmer, and K. Maier-Hein, "Brain tumor segmentation using large receptive field deep convolutional neural networks," in *Bildverarbeitung fur die Medizin 2017*, pp. 86–91, Kluwer Academic Publishers, 2017.

[182] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. S. Kirby, J. B. Freymann, K. Farahani, and C. Davatzikos, "Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features," *Sci Data*, vol. 4, p. 170117, 09 2017.

[183] P. Kleihues, W. Cavenee, *et al.*, "Who classification of tumours," *Pathology & genetics. Tumors of the nervous system. Lyon, France: IARCpress*, 2000.

[184] A. M. Rutman *et al.*, "Radiogenomics: creating a link between molecular diagnostics and diagnostic imaging," *European journal of radiology*, vol. 70, no. 2, 2009.

[185] R. Fukuma *et al.*, "Prediction of idh and tert promoter mutations in low-grade glioma from magnetic resonance images using a convolutional neural network," *Scientific Reports*, vol. 9, no. 1, 2019.

[186] K. Chang *et al.*, "Residual convolutional neural network for the determination of idh status in low-and high-grade gliomas from mr imaging," *Clinical Cancer Research*, vol. 24, no. 5, 2018.

[187] K. E. Emblem *et al.*, "Machine learning in preoperative glioma mri: Survival associations by perfusion-based support vector machine outperforms traditional mri," *Journal of magnetic resonance imaging*, vol. 40, no. 1, 2014.

[188] P. Kickingereder *et al.*, "Radiomic profiling of glioblastoma: identifying an imaging predictor of patient survival with improved performance over established clinical and radiologic risk models," *Radiology*, vol. 280, no. 3, 2016.

[189] P. Kickingereder *et al.*, "Radiomic subtyping improves disease stratification beyond key molecular, clinical, and standard imaging characteristics in patients with glioblastoma," *Neuro-oncology*, vol. 20, no. 6, 2018.

[190] X. Liu *et al.*, "A radiomic signature as a non-invasive predictor of progression-free survival in patients with lower-grade gliomas," *NeuroImage: Clinical*, vol. 20, 2018.

[191] Q. Li *et al.*, "A fully-automatic multiparametric radiomics model: towards reproducible and prognostic imaging signature for prediction of overall survival in glioblastoma multiforme," *Scientific reports*, vol. 7, no. 1, 2017.

[192] P. Kickingereder *et al.*, "Large-scale radiomic profiling of recurrent glioblastoma identifies an imaging predictor for stratifying anti-angiogenic treatment response," *Clinical Cancer Research*, vol. 22, no. 23, 2016.

[193] V. Frattini, V. Trifonov, J. M. Chan, A. Castano, M. Lia, F. Abate, S. T. Keir, A. X. Ji, P. Zoppoli, F. Niola, *et al.*, "The integrated landscape of driver genomic alterations in glioblastoma," *Nature genetics*, vol. 45, no. 10, pp. 1141–1149, 2013.

[194] J. Kim, I.-H. Lee, H. J. Cho, C.-K. Park, Y.-S. Jung, Y. Kim, S. H. Nam, B. S. Kim, M. D. Johnson, D.-S. Kong, *et al.*, "Spatiotemporal evolution of the primary glioblastoma genome," *Cancer cell*, vol. 28, no. 3, pp. 318–328, 2015.

[195] M. J. Ellis, M. Gillette, S. A. Carr, A. G. Paulovich, R. D. Smith, K. K. Rodland, R. R. Townsend, C. Kinsinger, M. Mesri, H. Rodriguez, *et al.*, "Connecting genomic alterations to cancer biology with proteomics: the nci clinical proteomic tumor analysis consortium," *Cancer discovery*, vol. 3, no. 10, pp. 1108–1112, 2013.

[196] N. Shah, X. Feng, M. Lankerovich, R. B. Puchalski, and B. Keogh, "Data from ivy gap," *The Cancer Imaging Archive*, vol. 10, p. K9, 2016.

[197] N. Pedano, A. E. Flanders, L. Scarpace, T. Mikkelsen, J. M. Eschbacher, B. Hermes, and Q. Ostrom, "Radiology data from the cancer genome atlas low grade glioma [tcga-lgg] collection," *Cancer Imaging Arch*, vol. 2, 2016.

[198] L. Scarpace, T. Mikkelsen, S. Cha, S. Rao, S. Tekchandani, D. Gutman, J. Saltz, B. Erickson, N. Pedano, A. Flanders, *et al.*, "Radiology data from the cancer genome atlas glioblastoma multiforme [tcga-gbm] collection. the cancer imaging archive," 2016.

[199] R. B. Puchalski, N. Shah, J. Miller, R. Dalley, S. R. Nomura, J.-G. Yoon, K. A. Smith, M. Lankerovich, D. Bertagnolli, K. Bickley, *et al.*, "An anatomic transcriptional atlas of human glioblastoma," *Science*, vol. 360, no. 6389, pp. 660–663, 2018.

[200] J. L. Boxerman, Z. Zhang, Y. Safriel, M. Larvie, B. S. Snyder, R. Jain, T. L. Chi, A. G. Sorensen, M. R. Gilbert, and D. P. Barboriak, "Early post-bevacizumab progression on contrast-enhanced mri as a prognostic marker for overall survival in recurrent glioblastoma: results from the acrin 6677/rtog 0625 central reader study," *Neuro-oncology*, vol. 15, no. 7, pp. 945–954, 2013.

[201] L. Scarpace, A. Flanders, R. Jain, T. Mikkelsen, and D. W. Andrews, "Data from rembrandt. the cancer imaging archive," 2015.

[202] J. Wang, E. Cazzato, E. Ladewig, V. Frattini, D. I. Rosenbloom, S. Zairis, F. Abate, Z. Liu, O. Elliott, Y.-J. Shin, *et al.*, "Clonal evolution of glioblastoma under therapy," *Nature genetics*, vol. 48, no. 7, pp. 768–776, 2016.

[203] C. Neftel, J. Laffy, M. G. Filbin, T. Hara, M. E. Shore, G. J. Rahme, A. R. Richman, D. Silverbush, M. L. Shaw, C. M. Hebert, *et al.*, "An integrative model of cellular states, plasticity, and genetics for glioblastoma," *Cell*, vol. 178, no. 4, pp. 835–849, 2019.

[204] N. D. Mathewson, O. Ashenberg, I. Tirosh, S. Gritsch, E. M. Perez, S. Marx, L. Jerby-Arnon, R. Chanoch-Myers, T. Hara, A. R. Richman, *et al.*, "Inhibitory cd161 receptor identified in glioma-infiltrating t cells by single-cell analysis," *Cell*, vol. 184, no. 5, pp. 1281–1298, 2021.

[205] H. Kim, S. Zheng, S. S. Amini, S. M. Virk, T. Mikkelsen, D. J. Brat, J. Grimsby, C. Sougnez, F. Muller, J. Hu, *et al.*, "Whole-genome and multisector exome sequencing of primary and post-treatment glioblastoma reveals patterns of tumor evolution," *Genome research*, vol. 25, no. 3, pp. 316–327, 2015.

[206] T. Mazor, A. Pankov, B. E. Johnson, C. Hong, E. G. Hamilton, R. J. Bell, I. V. Smirnov, G. F. Reis, J. J. Phillips, M. J. Barnes, *et al.*, "Dna methylation and somatic mutations converge on the cell cycle and define similar evolutionary histories in brain tumors," *Cancer cell*, vol. 28, no. 3, pp. 307–317, 2015.

[207] Q. Wang, B. Hu, X. Hu, H. Kim, M. Squatrito, L. Scarpace, A. C. deCarvalho, S. Lyu, P. Li, Y. Li, *et al.*, "Tumor evolution of glioma-intrinsic gene expression subtypes associates with immunological changes in the microenvironment," *Cancer cell*, vol. 32, no. 1, pp. 42–56, 2017.

[208] F. Nassiri, A. Chakravarthy, S. Feng, S. Y. Shen, R. Nejad, J. A. Zuccato, M. R. Voisin, V. Patil, C. Horbinski, K. Aldape, *et al.*, "Detection and discrimination of intracranial tumors using plasma cell-free dna methylomes," *Nature Medicine*, vol. 26, no. 7, pp. 1044–1047, 2020.

[209] M. E. Hegi, A.-C. Diserens, T. Gorlia, M.-F. Hamou, N. De Tribolet, M. Weller, J. M. Kros, J. A. Hainfellner, W. Mason, L. Mariani, *et al.*, "Mgmt gene silencing and benefit from temozolomide in glioblastoma," *New England Journal of Medicine*, vol. 352, no. 10, pp. 997–1003, 2005.

[210] W. Yu, L. Zhang, Q. Wei, and A. Shao, "O6-methylguanine-dna methyltransferase (mgmt): challenges and new opportunities in glioma chemotherapy," *Frontiers in oncology*, vol. 9, p. 1547, 2020.

[211] R. J. Molenaar, D. Verbaan, S. Lamba, C. Zanon, J. W. Jeuken, S. H. Boots-Sprenger, P. Wesseling, T. J. Hulsebos, D. Troost, A. A. Van Tilborg, *et al.*, "The combination of idh1 mutations and mgmt methylation status predicts survival in glioblastoma better than either idh1 or mgmt alone," *Neuro-oncology*, vol. 16, no. 9, pp. 1263–1273, 2014.

[212] R. J. Leary, M. Sausen, I. Kinde, N. Papadopoulos, J. D. Carpten, D. Craig, J. O'Shaughnessy, K. W. Kinzler, G. Parmigiani, B. Vogelstein, *et al.*, "Detection of chromosomal alterations in the circulation of cancer patients with whole-genome sequencing," *Science translational medicine*, vol. 4, no. 162, pp. 162ra154–162ra154, 2012.

[213] D.-W. Shin, S. Lee, S. W. Song, Y. H. Cho, S. H. Hong, J. H. Kim, H. S. Kim, J. E. Park, S. J. Nam, and Y.-H. Kim, "Survival outcome and prognostic factors in anaplastic oligodendroglioma: a single-institution study of 95 cases," *Scientific Reports*, vol. 10, no. 1, pp. 1–8, 2020.

[214] X. Hu, E. Martinez-Ledesma, S. Zheng, H. Kim, F. Barthel, T. Jiang, K. R. Hess, and R. G. Verhaak, "Multigene signature for predicting prognosis of patients with 1p19q co-deletion diffuse glioma," *Neuro-oncology*, vol. 19, no. 6, pp. 786–795, 2017.

[215] C. J. Kinslow, A. L. Garton, A. I. Rae, L. P. Marcus, C. M. Adams, G. M. McKhann, M. B. Sisti, E. S. Connolly, J. N. Bruce, A. I. Neugut, *et al.*, "Extent of resection and survival for oligodendroglioma: a us population-based study," *Journal of Neuro-Oncology*, vol. 144, no. 3, pp. 591–601, 2019.

[216] T. E. Richardson, M. Williams, K. Galbraith, K. Mirchia, A. Kumar, C. Xing, and J. M. Walker, "Total copy number variation, somatic mutation burden, and histologic grade correlate with clinical outcome in oligodendroglioma," *Clinical neuropathology*, vol. 39, no. 5, pp. 238–242, 2020.

[217] W. Yan, W. Zhang, G. You, J. Zhang, L. Han, Z. Bao, Y. Wang, Y. Liu, C. Jiang, C. Kang, *et al.*, "Molecular classification of gliomas based on whole genome gene expression: a systematic report of 225 samples from the chinese glioma cooperative group," *Neuro-oncology*, vol. 14, no. 12, pp. 1432–1440, 2012.

[218] S. Madhavan, J.-C. Zenklusen, Y. Kotliarov, H. Sahni, H. A. Fine, and K. Buetow, "Rembrandt: helping personalized medicine become a reality through integrative translational research," *Molecular cancer research*, vol. 7, no. 2, pp. 157–167, 2009.

[219] L. A. Gravendeel, M. C. Kouwenhoven, O. Gevaert, J. J. de Rooi, A. P. Stubbs, J. E. Duijm, A. Daemen, F. E. Bleeker, L. B. Bralten, N. K. Kloosterhof, *et al.*, "Intrinsic gene expression profiles of gliomas are a better predictor of survival than histology," *Cancer research*, vol. 69, no. 23, pp. 9065–9072, 2009.

[220] Z.-S. Bao, H.-M. Chen, M.-Y. Yang, C.-B. Zhang, K. Yu, W.-L. Ye, B.-Q. Hu, W. Yan, W. Zhang, J. Akers, *et al.*, "Rna-seq of 272 gliomas revealed a novel, recurrent ptprz1-met fusion transcript in secondary glioblastomas," *Genome research*, vol. 24, no. 11, pp. 1765–1773, 2014.

[221] M. H. Cohen, J. R. Johnson, and R. Pazdur, "Food and drug administration drug approval summary: temozolomide plus radiation therapy for the treatment of newly diagnosed glioblastoma multiforme," *Clinical Cancer Research*, vol. 11, no. 19, pp. 6767–6771, 2005.

[222] D. W. Parsons, S. Jones, X. Zhang, J. C.-H. Lin, R. J. Leary, P. Angenendt, P. Mankoo, H. Carter, I.-M. Siu, G. L. Gallia, *et al.*, "An integrated genomic analysis of human glioblastoma multiforme," *science*, vol. 321, no. 5897, pp. 1807–1812, 2008.

[223] R. Stupp, S. Taillibert, A. Kanner, W. Read, D. M. Steinberg, B. Lhermitte, S. Toms, A. Idbaih, M. S. Ahluwalia, K. Fink, *et al.*, "Effect of tumor-treating fields plus maintenance temozolomide vs maintenance temozolomide alone on survival in patients with glioblastoma: a randomized clinical trial," *Jama*, vol. 318, no. 23, pp. 2306–2316, 2017.

[224] D. Capper, D. Stichel, F. Sahm, D. T. Jones, D. Schrimpf, M. Sill, S. Schmid, V. Hovestadt, D. E. Reuss, C. Koelsche, *et al.*, "Practical implementation of dna methylation and copy-number-based cns tumor diagnostics: the heidelberg experience," *Acta neuropathologica*, vol. 136, no. 2, pp. 181–210, 2018.

[225] J.-P. Fortin, T. J. Triche Jr, and K. D. Hansen, "Preprocessing, normalization and integration of the illumina humanmethylationepic array with minfi," *Bioinformatics*, vol. 33, no. 4, pp. 558–560, 2017.

[226] V. Hovestadt and M. Zapatka, "Enhanced copy-number variation analysis using illumina dna methylation arrays," *R package version*, vol. 1, no. 0, 2019.

[227] S. Xu, L. Tang, X. Li, F. Fan, and Z. Liu, "Immunotherapy for glioma: Current management and future application," *Cancer letters*, vol. 476, pp. 1–12, 2020.

[228] D. A. Reardon, A. A. Brandes, A. Omuro, P. Mulholland, M. Lim, A. Wick, J. Baehring, M. S. Ahluwalia, P. Roth, O. Bähr, *et al.*, "Effect of nivolumab vs bevacizumab in patients with recurrent glioblastoma: the checkmate 143 phase 3 randomized clinical trial," *JAMA oncology*, vol. 6, no. 7, pp. 1003–1010, 2020.

[229] J. R. Fink, M. Muzi, M. Peck, and K. A. Krohn, "Multimodality brain tumor imaging: Mr imaging, pet, and pet/mr imaging," *Journal of Nuclear Medicine*, vol. 56, no. 10, pp. 1554–1561, 2015.

Appendix A

SUPPLEMENTAL FIGURES

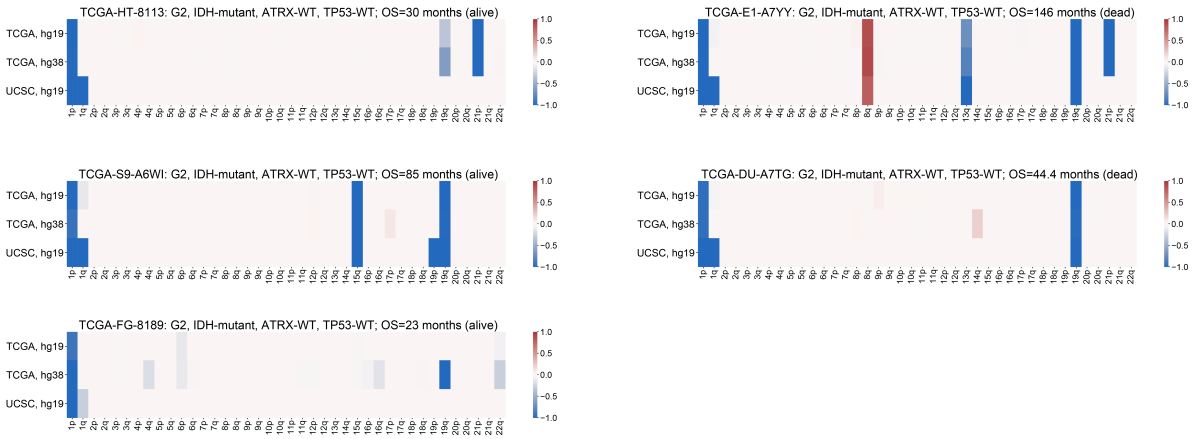


Figure A.1: SCNA Profiles of excluded TCGA IDH-mutant tumors. In four of the five TCGA IDH-mutant samples we excluded from our study, the UCSC hg19 pipeline called monosomy of chromosome 1 while the other pipelines called 1q intact. In one case, the GDC hg38 pipeline called loss of 19q, but the other two pipelines did not. None of these patients harbored *TP53* or *ATRX* mutations, common in IDH-mutant astrocytomas, and thus we suspect they are oligodendrogliomas despite their copy number profile not consistently support that claim.

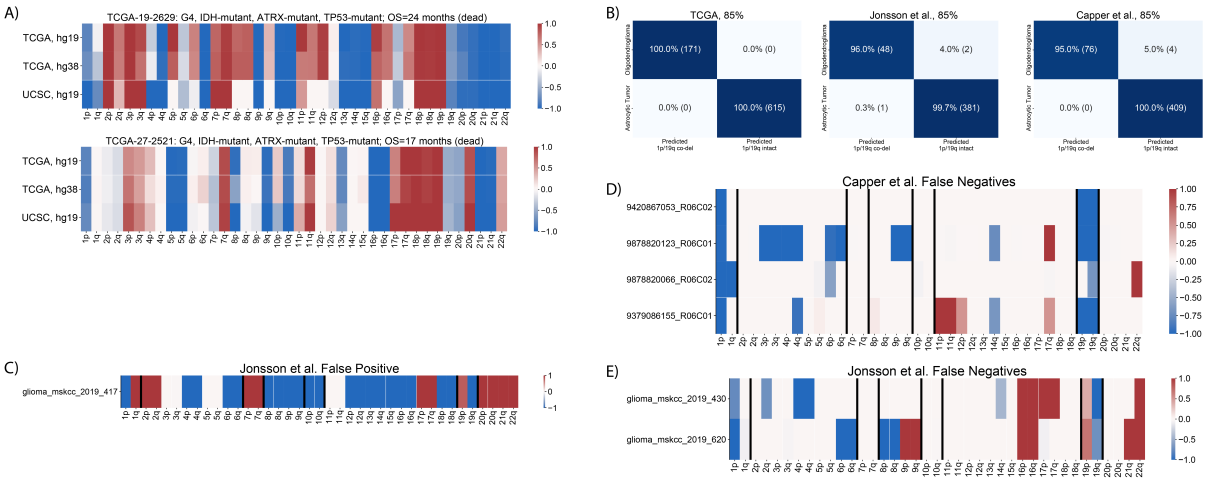


Figure A.2: Development and validation of 1p/19q-codeletion threshold. A) SCNA profiles of two TCGA grade 4 IDH-mutant astrocytomas that would have been considered 1p/19q-codeleted at thresholds lower than 85% (75% and 80%). B) Confusion matrixes show our 1p/19q-codeletion screen on the TCGA training set and validation sets published by Jonsson et al. and Capper et al. were accurate. C) Of the 921 astrocytic tumors in the validation sets published by Jonsson et al. and Capper et al., our 1p/19q co-deletion screen produced only 1 false positive. This patient is a grade 4 IDH-mutant astrocytoma that harbors high copy number alteration burden, including loss of 1p and 19q, co-gain of 1q and 19p, and +7/-10. We recommend examining the SCNA profiles of SCNA-burdened tumors predicted to be 1p/19q-codeleted. D) Four of the six false negative tumors from our 1p/19q-codeletion validation screen are from the dataset published by Capper et al. SCNA data indicates that each of these tumors harbored monosomy of chromosome 1 and/or monosomy of chromosome 19, and therefore are not 1p/19q-codeleted. Thus, these tumors do not represent model errors—they represent either misdiagnosed patients or SCNA pipeline errors. E) The remaining two false negative tumors from our 1p/19q-codeletion screen validation fail to meet our 85% threshold: one meets a threshold of 75% and the other meets a threshold of 71%.

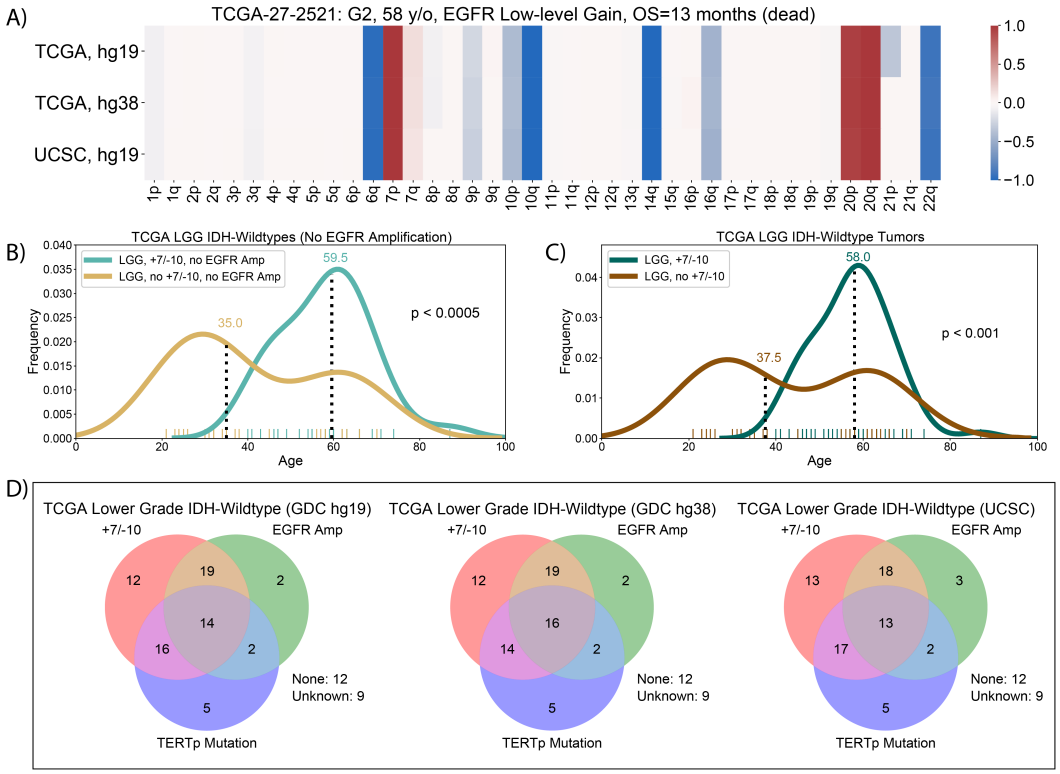


Figure A.3: SCNA profiles, age associations, and training set inclusion implications of our +7/-10 threshold. A) The SCNA profile of the TCGA histological lower-grade IDH-wildtype glioma without EGFR amplification closest to our 50% threshold for +7/-10. B) Among TCGA histological lower grade IDH-wildtype glioma without EGFR amplification, those with +7/-10 are significantly older than those without +7/-10. C) Among TCGA histological lower-grade IDH-wildtype glioma with or without EGFR amplification, those with +7/-10 are significantly older than those without +7/-10. D) The SCNA profile of the TCGA histological lower grade IDH-wildtype glioma with no TERTp mutation, no +7/-10, and with inconsistent EGFR amplification status across three SCNA pipelines. E) The classification of TCGA histological lower-grade IDH-wildtype glioma as pediatric-type or IDH-wildtype glioblastoma varies based on SCNA pipeline choice. The source of this variation is EGFR amplification calling inconsistency.

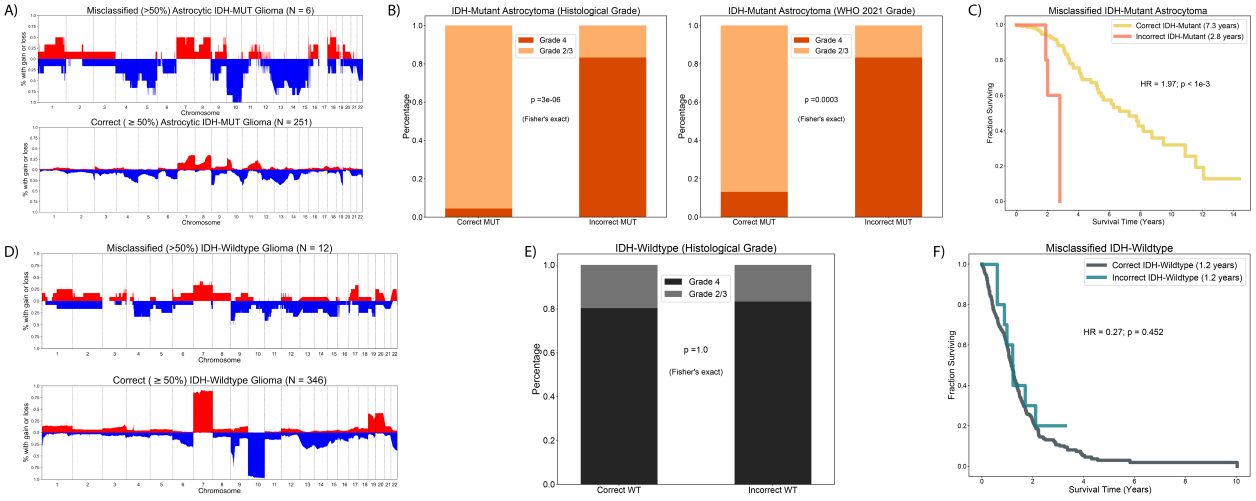


Figure A.4: Differences between correct and incorrect IDH mutation predictions on TCGA astrocytic tumors during cross-validation. A) Misclassified IDH-mutant astrocytomas had increased copy number burden, especially on chromosome 10. B) Misclassified IDH-mutant astrocytomas were disproportionately grade 4 before and after lower grade IDH-mutant astrocytomas with CDKN2A homozygous deletion were promoted to grade 4 in accordance with the guidelines proposed for the fifth edition of the WHO classification of CNS tumors. C) Misclassified IDH-mutant astrocytomas followed a clinical course significantly worse than correctly classified IDH-mutant astrocytomas. D) Misclassified IDH-wildtype glioblastomas tended to have fewer SCNAs than their correctly classified counterparts, especially on chromosomes 7 and 10. E) No difference in histological grade between correctly and incorrectly classified IDH-wildtype glioblastomas was observed. F) No difference in patient outcome between correctly and incorrectly classified IDH-wildtype glioblastomas was observed.

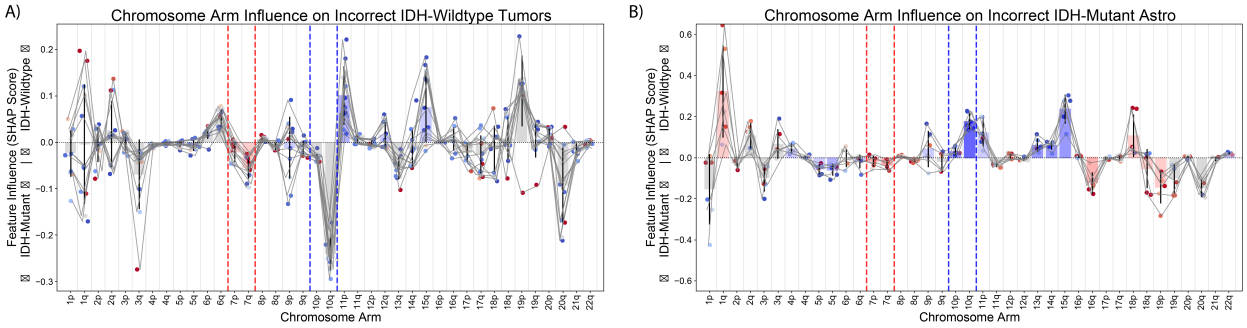


Figure A.5: Atypical chromosome arm 10q SCNA status drives IDH mutation classifier mistakes during cross-validation. A) A SHAP interpretation of the IDH mutation classifier's mistakes on IDH-wildtype tumors indicates that intact 10q status along with low-level gains of on 7q drove incorrect predictions. Lines are drawn connecting points associated with the same patient. Each incorrectly predicted patient is assigned a point per chromosome arm. The magnitude of each point's vertical coordinate indicates how influential the chromosome arm was to the classifier's prediction. Positive vertical values favor IDH-wildtype predictions; negative values indicate a preference for IDH-mutant astrocytoma predictions. The color of each point corresponds to the average chromosome arm SCNA value: blue indicates loss, and red indicates gain. The magnitude of the bars drawn for each chromosome arm is the average chromosome arm SHAP score for all patients plotted; the bar's color indicates the average SCNA state across all patients plotted. B) A SHAP interpretation of the IDH mutation classifier's mistakes on IDH-mutant astrocytomas indicates that losses on 10q and 15q and gains on chromosome arm 1q drove incorrect predictions.

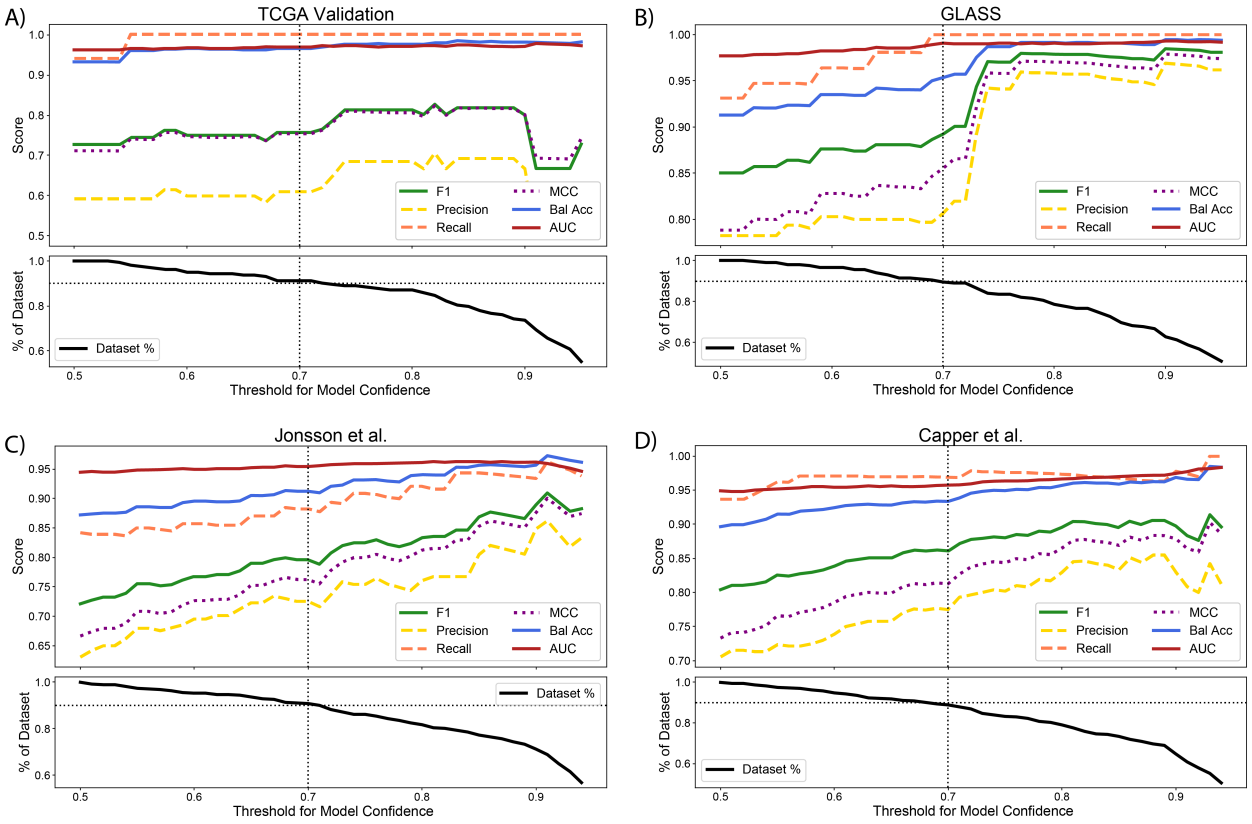


Figure A.6: Model calibration increases the performance of the astrocytic IDH-mutation classifier across four validation sets. Model calibration shows that the performance of our astrocytic IDH-mutation classifier increases with prediction confidence in four validation sets: the holdout TCGA validation set (A), the GLASS dataset (B), and datasets published Jonsson et al. (C) and Capper et al. (D).

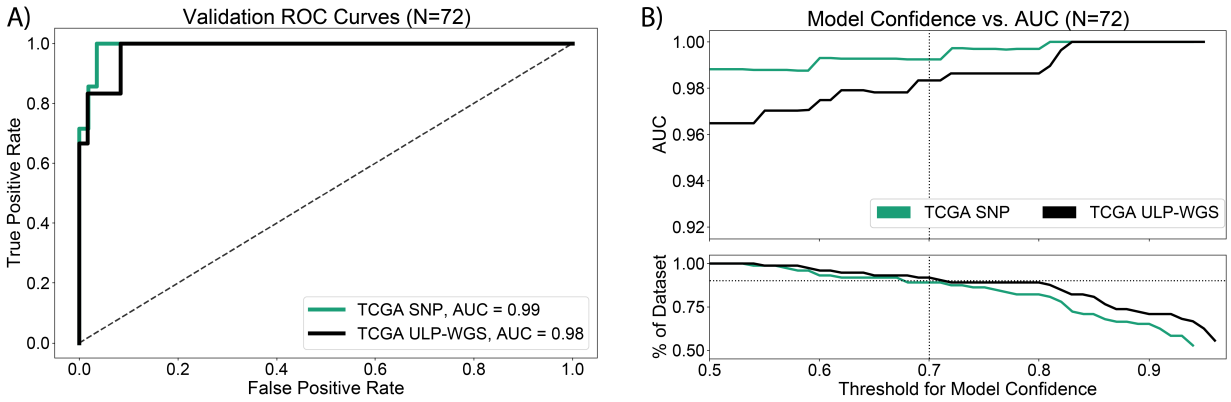


Figure A.7: Validation results for the IDH mutation classifier for TCGA SNP array derived and ultra-low-pass whole genome sequencing (ULP-WGS) derived SCNA data. A) IDH mutation classifier ROC curves for TCGA patients with computed ULP-WGS derived SCNA data. The high confidence predictions (confidence ¿ 70%) based on SNP derived and ULP-WGS derived SCNA data were very similar: one astrocytic tumor with ULP-WGS-like data in the TCGA validation set was predicted differently (ULP-WGS AUC=0.98, SNP AUC=0.99). B) Furthermore, disregarding lower-confidence predictions (confidence ¡ 70%) made from ULP-WGS derived SCNA data does not exclude more patients than from SNP derived SCNA data (10% for each).

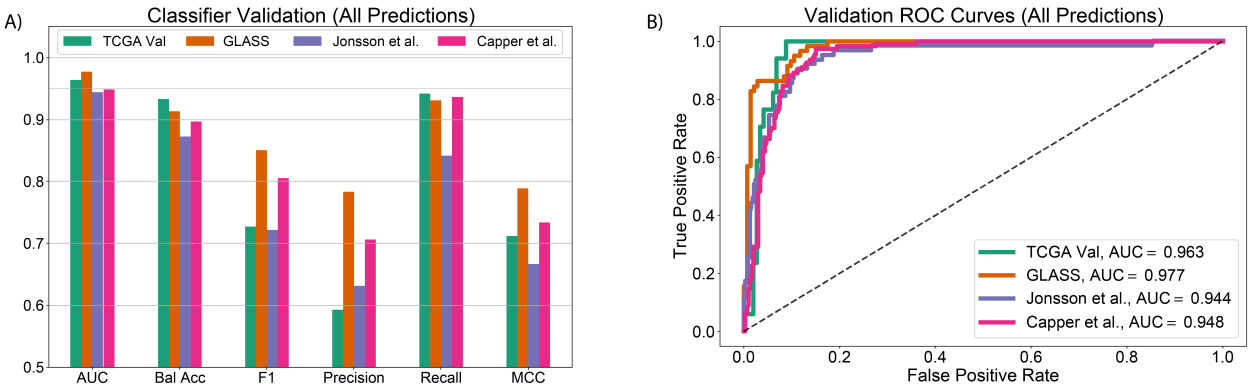


Figure A.8: Validation results for the IDH mutation classifier for all samples across four validation sets. A) IDH mutation classifier results for predictions over all patients and validation sets. The model performed well across all metrics except for precision, indicating that IDH-wildtype tumors were more difficult to identify than IDH-mutation astrocytomas. B) AUC scores were slightly lower than those achieved by evaluating the IDH-mutation classifier on predictions with 70% confidence or greater, but they were still above 0.94 on all four validation sets.

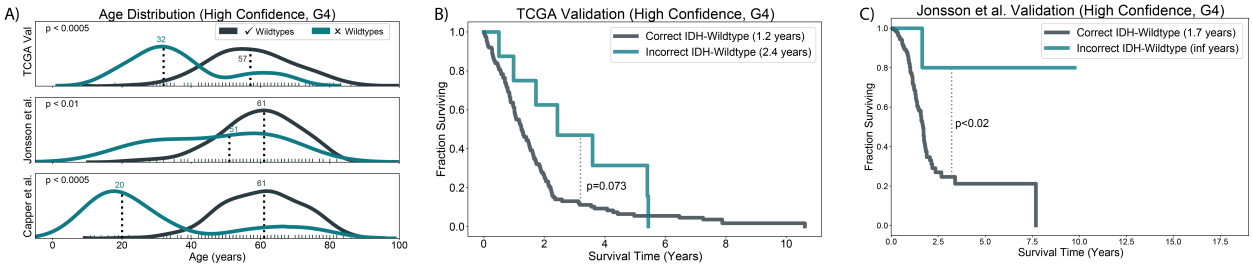


Figure A.9: Misclassified histological grade 4 IDH-wildtype tumors in our validation set were younger and survived longer than their correctly classified counterparts. A) When our age and survival analysis is restricted to histological grade 4 IDH-wildtype glioblastomas, tumors misclassified by our IDH-mutation classifier remained significantly younger than correctly classified tumors across three validation sets. B) Because our TCGA validation set is entirely composed of grade 4 tumors, misclassified IDH-wildtype tumors in this dataset survive better than correctly classified IDH-wildtype tumors with the same marginal significance reported in the main text. C) On the dataset published by Jonsson et al., correctly misclassified histological grade 4 IDH-wildtype tumors survived significantly better than correctly classified histological grade 4 IDH-wildtype tumors.

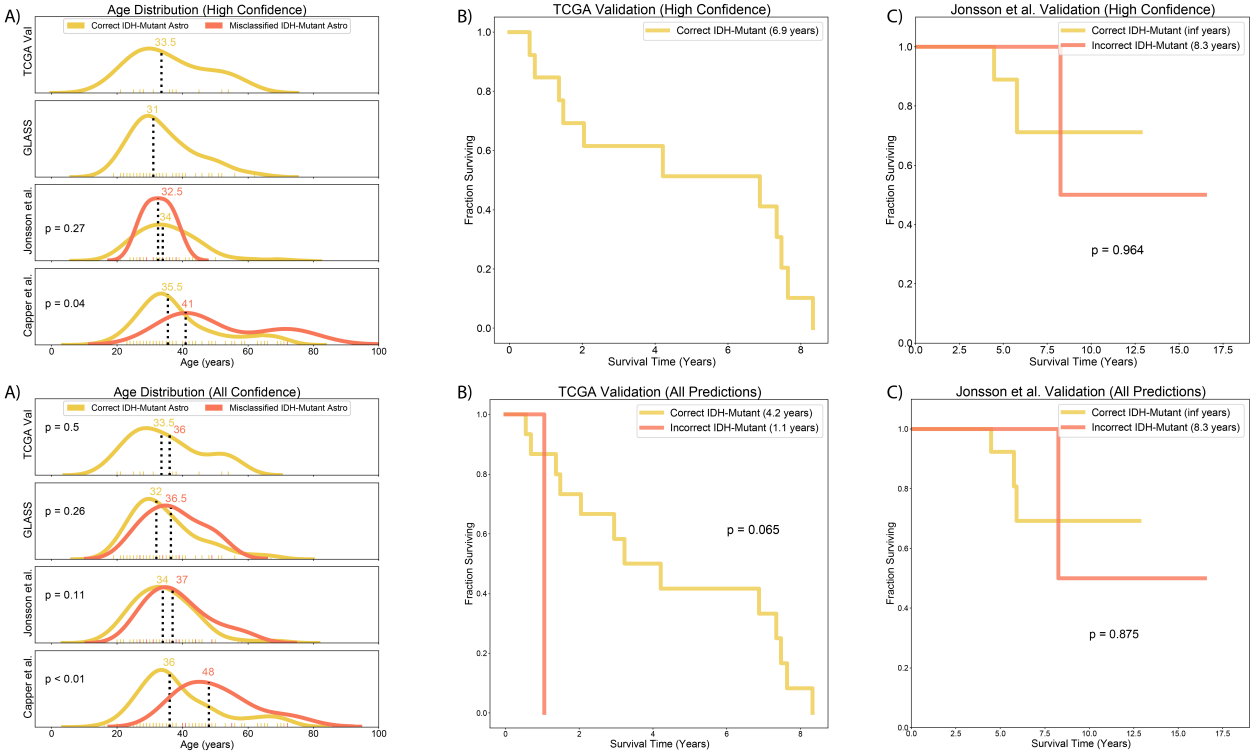


Figure A.10: No consistent age differences between correctly and incorrectly classified IDH-mutant astrocytomas were observed. A) Among high-confidence predictions, we did not observe a consistent, significant age difference between correctly and incorrectly predicted IDH-mutant astrocytomas over three independent validation sets and the holdout TCGA validation set. B, C) Among high-confidence validation predictions in our TCGA holdout validation set and a dataset published by Jonsson et al., we did not observe a significant difference in patient outcome between correctly and incorrectly predicted IDH-mutant astrocytomas. D) Overall patient predictions, we did not observe a consistent, significant age difference between correctly and incorrectly predicted IDH-mutant astrocytomas over three independent validation sets and the holdout TCGA validation set. E, F) Overall patient predictions in our TCGA holdout validation set and a dataset published by Jonsson et al., we did not observe a significant difference in patient outcome between correctly and incorrectly predicted IDH-mutant astrocytomas.

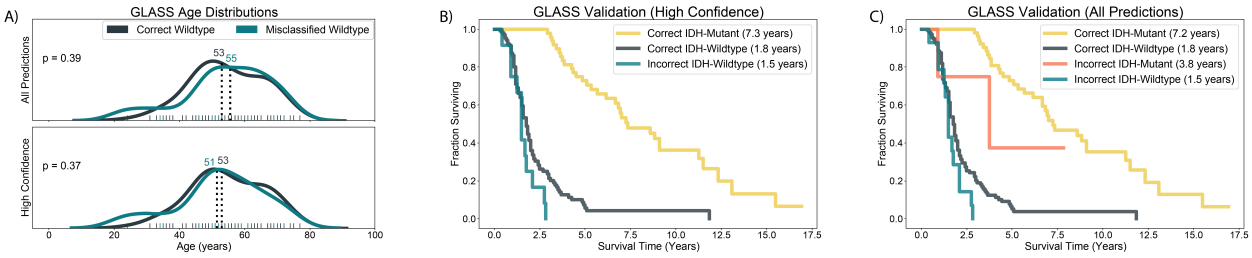


Figure A.11: Patients in the GLASS dataset do not follow the same age or survival patterns found in the other validation datasets. A) No significant age difference between correct and incorrectly classified IDH-wildtype glioblastomas was observed in the GLASS dataset. B, C) The tendency for misclassified IDH-wildtype glioblastomas to outlive correctly classified IDH-wildtype glioblastomas was not observed in the GLASS dataset.

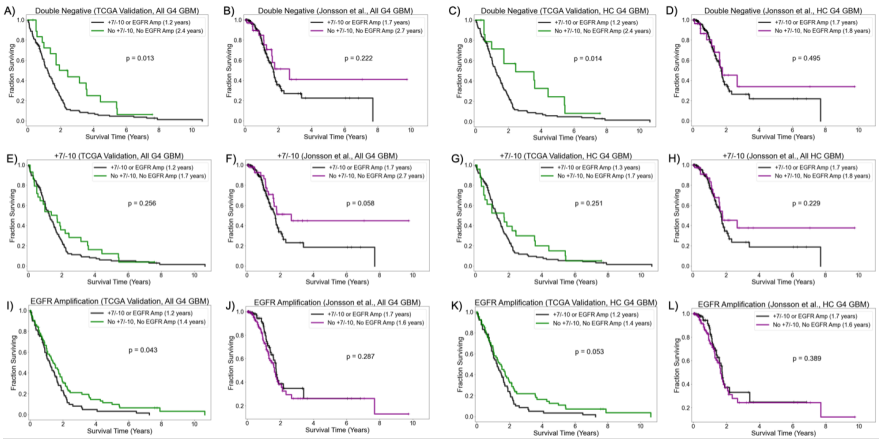


Figure A.12: Survival differences incurred by +7/-10 and EGFR amp in the TCGA validation set and the Jonsson et al. validation set are inconsistent among histological grade 4 (hG4) IDH-wildtype GBM when examining all such tumors (All) and when examining only those predicted by our model with high confidence (HC). A-B) Among all available hG4 IDH-wildtype GBMs in the TCGA validation and the Jonsson et al. dataset, double negative patient (no +7/-10, no EGFR amp) do significantly better in the TCGA validation dataset than patients with either +7/-10 or EGFR amp, unlike similar patients in the Jonsson et al. validation dataset. C-D) Among hG4 IDH-wildtype GBMs in the TCGA validation and the Jonsson et al. dataset that our model predicted with HC, double negative patient do significantly better in the TCGA validation dataset than patients with either +7/-10 or EGFR amp; this is not true for similar patients in the Jonsson et al. validation dataset. E-F) Among all available hG4 IDH-wildtype GBMs in the TCGA validation and the Jonsson et al. dataset, patients lacking +7/-10 do not do significantly better in the TCGA validation dataset than patients with +7/-10; in the Jonsson et al. dataset, patients lacking +7/-10 have marginally significantly better outcomes. G-H) Among hG4 IDH-wildtype GBMs in the TCGA validation and the Jonsson et al. dataset that our model predicted with HC, patients lacking +7/-10 do not do significantly better in the TCGA validation dataset or Jonsson et al. validation dataset. I-J) Among all available hG4 IDH-wildtype GBMs in the TCGA validation and the Jonsson et al. dataset, patients lacking EGFR amp do better in the TCGA validation dataset but not in the Jonsson et al. dataset. K-L) Among hG4 IDH-wildtype GBMs in the TCGA validation and the Jonsson et al. dataset that our model predicted with HC, patients lacking EGFR amp do not do marginally significantly better in the TCGA validation dataset but not in the Jonsson et al. validation dataset.

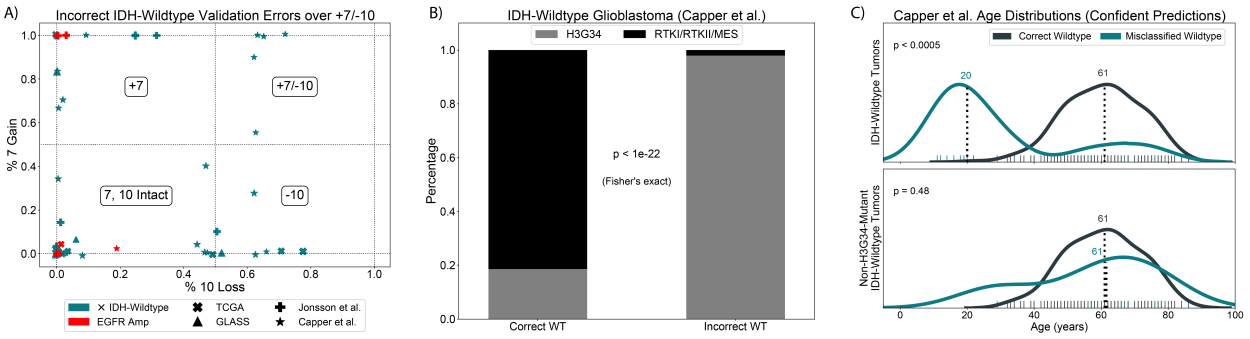


Figure A.13: Distribution of IDH-wildtype glioblastoma SCNA features and H3 G34-mutant gliomas in misclassified IDH-wildtype tumors. A) Of all misclassified histological WHO grade 4 IDH-wildtype gliomas in our three independent validation sets and the holdout TCGA validation set, only 19% (11/59) show molecular IDH-wildtype glioblastoma features +7/-10 or EGFR amplification. Point size indicates prediction confidence. B) In the dataset published by Capper et al., 88% of misclassified IDH-wildtype samples were WHO grade 4 diffuse hemispheric glioma, H3 G34-mutant, while only 2% of correctly classified samples were diffuse hemispheric glioma, H3 G34-mutant. C) The presence of diffuse hemispheric glioma, H3 G34-mutant in the dataset published by Capper et al. explained the age difference between correctly and incorrectly classified IDH-wildtype glioblastomas.

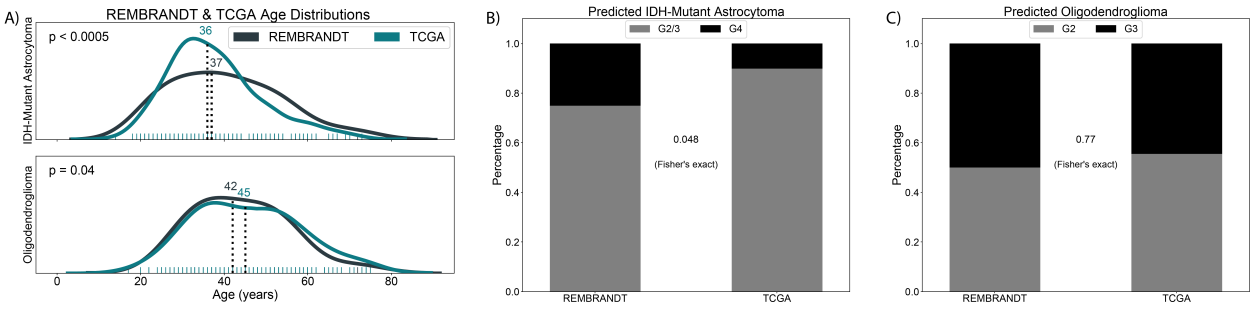


Figure A.14: Predicted REMBRANDT IDH-mutant astrocytomas harbor more risk factors than their TCGA counterparts, but the same is not true for predicted 1p/19q-codeleted oligodendroglioma. A) Predicted IDH-mutant astrocytomas in the REMBRANDT dataset are significantly older than ground truth IDH-mutant astrocytomas in our TCGA training set. Predicted 1p/19q-codeleted oligodendrogliomas are younger than ground truth 1p/19q-codeleted oligodendrogliomas in our TCGA training set, but the difference in median age is only 3 years. B) Predicted IDH-mutant astrocytomas in the REMBRANDT dataset are disproportionally grade 4 compared to ground truth IDH-mutant astrocytomas in our TCGA training set. C) Predicted 1p/19q-codeleted oligodendrogliomas and ground truth 1p/19q-codeleted oligodendrogliomas in our TCGA training set have a similar proportion of grade 2 and grade 3 tumors.

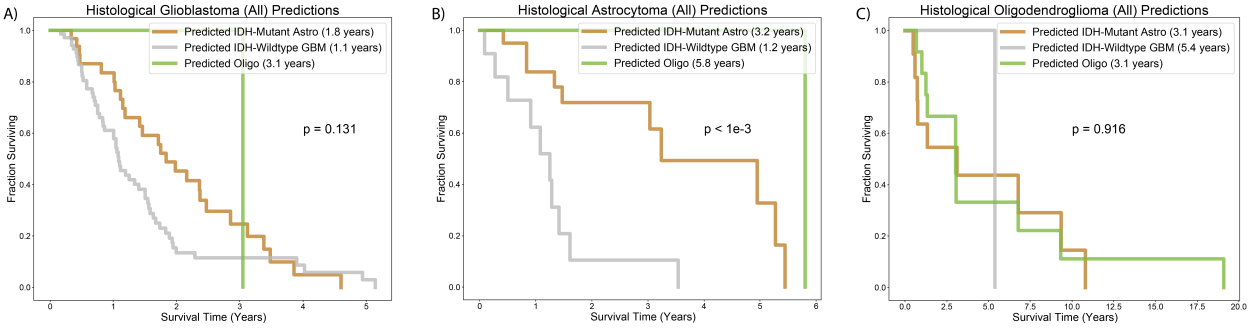


Figure A.15: Survival results for model predictions on all REMBRANDT patients. Kaplan–Meier curves are given for predicted adult diffuse glioma molecular subtype diagnosis of REMBRANDT patients with histological diagnosis of glioblastoma (A), astrocytoma (B), and oligodendroglioma (C).

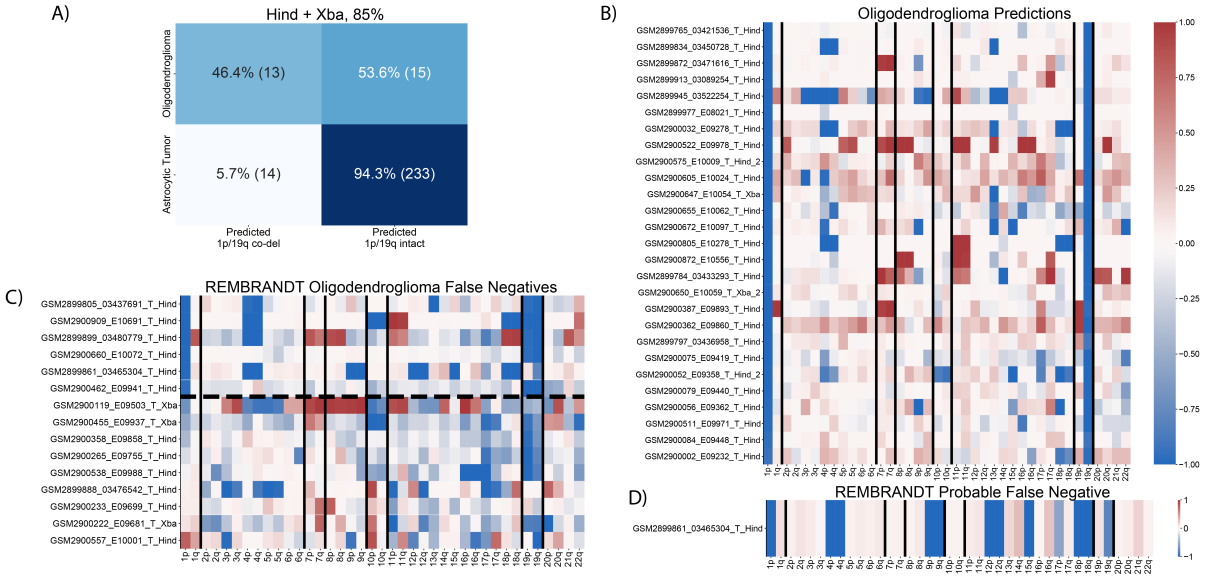


Figure A.16: Our 1p/19q-codeletion predictions were accurate on the REMBRANDT dataset despite discordance with histological diagnosis. A) Only 46% of REMBRANDT histological oligodendrogliomas were predicted to harbor 1p/19q co-deletions, and only 48% of REMBRADNT tumors harboring predicted 1p/19q co-deletions (n=27) were diagnosed as histological oligodendroglioma (n=13). B) Copy number profiles of predicted oligodendroglioma revealed clear 1p/19q co-deletions without loss of 1q or 19p. C) Of patients diagnosed with histological oligodendroglioma for whom we did not predict 1p/19q co-deletions, we saw that most (5/6, above dotted line) tumors that lose 1p harbor monosomy chromosome 19, prohibiting a 1p/19q-codeletion. D) SCNA profile of a likely misclassified 1p/19q-codeleted oligodendroglioma. This patient's proportion of 19q loss falls slightly below our 85% threshold (79%).

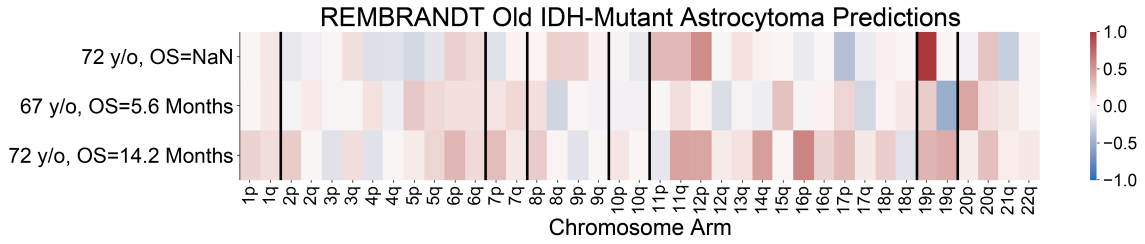


Figure A.17: The oldest predicted REMBRANDT IDH-mutant astrocytomas do not have IDH-wildtype glioblastoma-like SCNA profiles. Pictured here are SCNA profiles of four older (age range 65-80) patients diagnosed with predicted IDH-mutant astrocytoma. None harbor +7/-10 and only one shows EGFR amplification (72 y/o, G4, OS=20.6 Months).

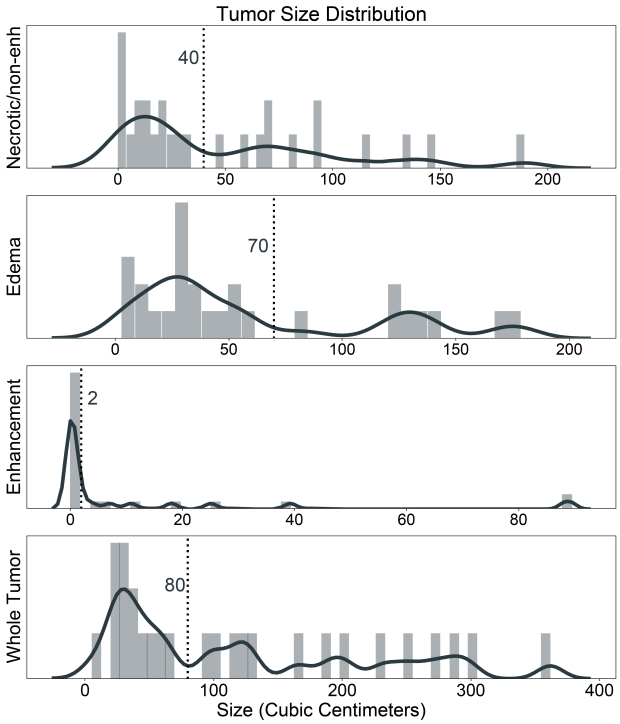


Figure A.18: MRI tumor compartment thresholds were determined visually. The volume of oligodendroglioma tumor compartments was binarized using visually determined thresholds. These thresholds were used to classify the necrotic/non-enhancing compartment, tumor edema compartment, and whole tumor volume as small or large. Tumor enhancement was classified as present (> 2 cm^3) or absent (< 2 cm^3).

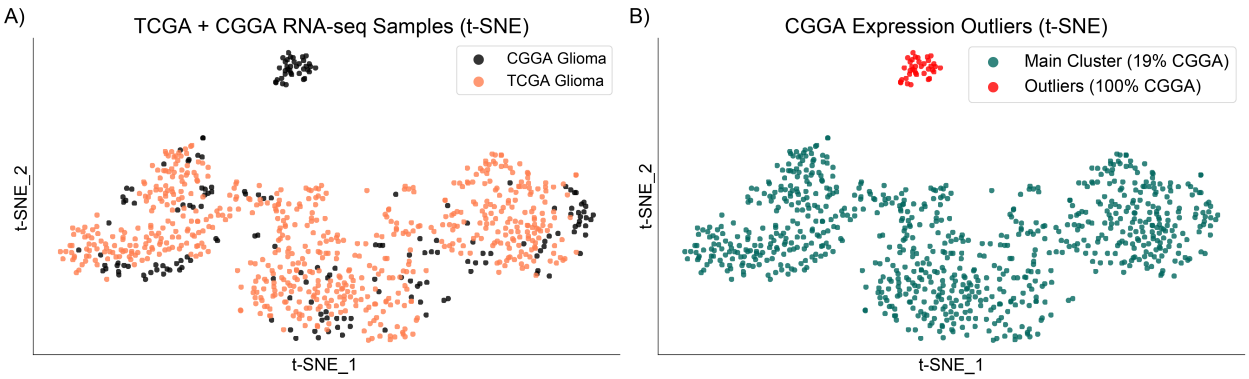


Figure A.19: CGGA RNA-seq data contains outliers. A) Unsupervised dimension reduction (t-SNE) of TCGA and CGGA RNA-seq TPM data computed from recount2 data shows that the TCGA and CGGA projection forms two distinct clusters. B) The smaller cluster consists of only CGGA patients, of which four are oligodendroglioma, IDH-mutant and 1p/19q-codeleted. We deemed these four tumors to be outliers because they appear dissimilar to all TCGA samples. It is unlikely that this cluster is the product of bioinformatic processes as all TCGA and CGGA TPM values were calculated using exactly the same pipeline.

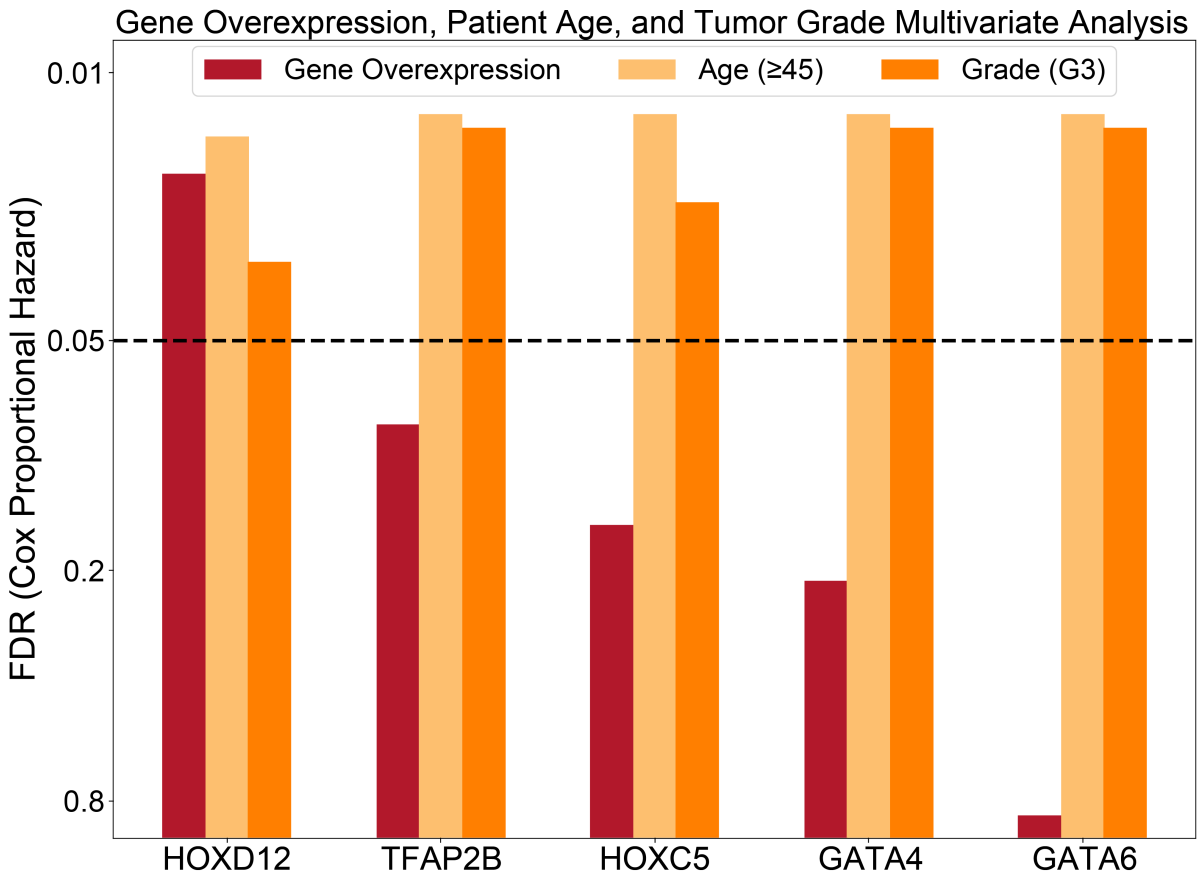


Figure A.20: *HOXD12* is the only age-associated gene whose overexpression is independently prognostic of oligodendroglioma age and tumor grade. Of the five age-associated genes whose overexpression is significantly prognostic in TCGA oligodendroglioma univariate tests, *HOXD12* is the only gene whose overexpression is significantly prognostic independent of age and grade after p-value correction (FDR).

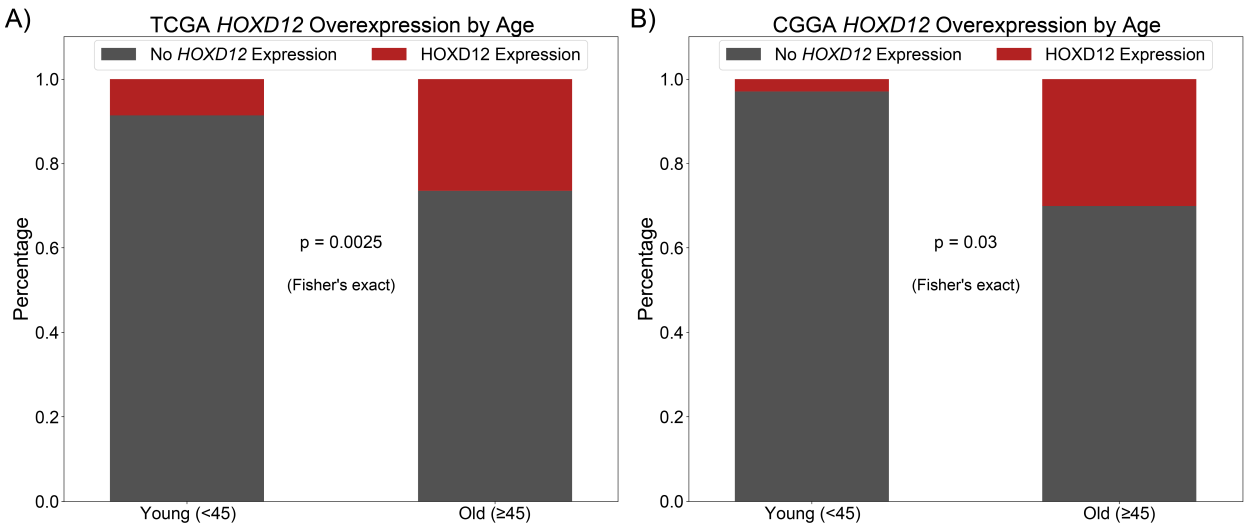


Figure A.21: *HOXD12* overexpression is age-associated in the TCGA and CGGA. A,B) Binning oligodendroglioma patients by TCGA median age (45 years old) and *HOXD12* overexpression (expressed or not), our observation that *HOXD12* overexpression was significantly associated with age in the TCGA is validated in the CGGA. TCGA age significance was tested using Wald test in the main text; here we binarize our data and use Fisher's exact test.

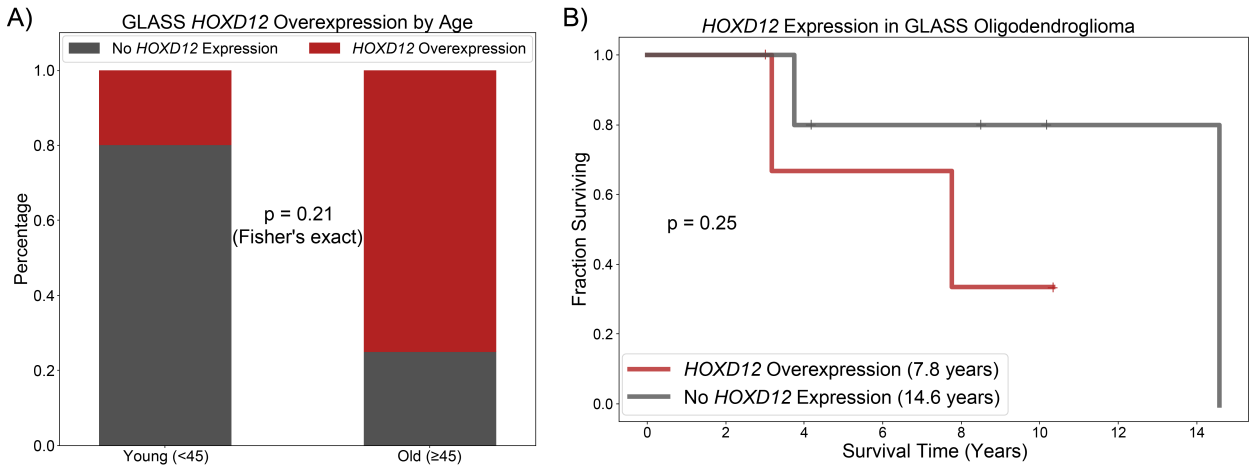


Figure A.22: *HOXD12* overexpression shows trends towards age and survival association in the GLASS dataset. A) Even though there are only 9 GLASS oligodendroglioma patients with RNA-seq data, 3 of the 4 older patients harbor *HOXD12* overexpression while only 1 of 5 younger patients harbor *HOXD12* overexpression. B) Similarly, despite the small size of the GLASS dataset, patients with *HOXD12* overexpression trend toward worse survival.

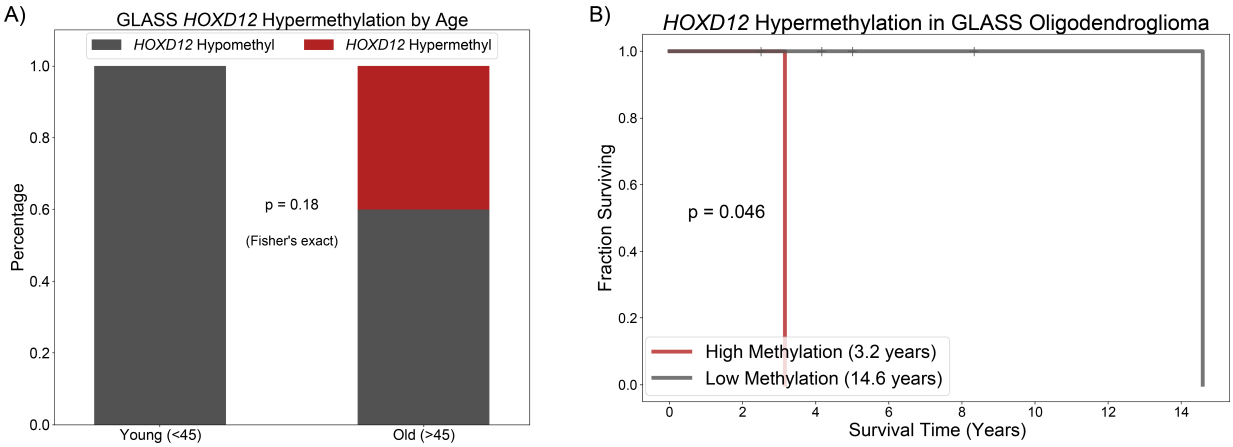


Figure A.23: *HOXD12* hypermethylation may be survival associated in the GLASS dataset. A) None of the 6 younger (¡ 45 years old) GLASS oligodendroglioma patients with DNA methylation data harbored *HOXD12* hypermethylation, while 2 of the 5 older GLASS patients did. B) Despite an apparent lack of age-association, GLASS patients with *HOXD12* hypermethylation have significantly worse outcomes.

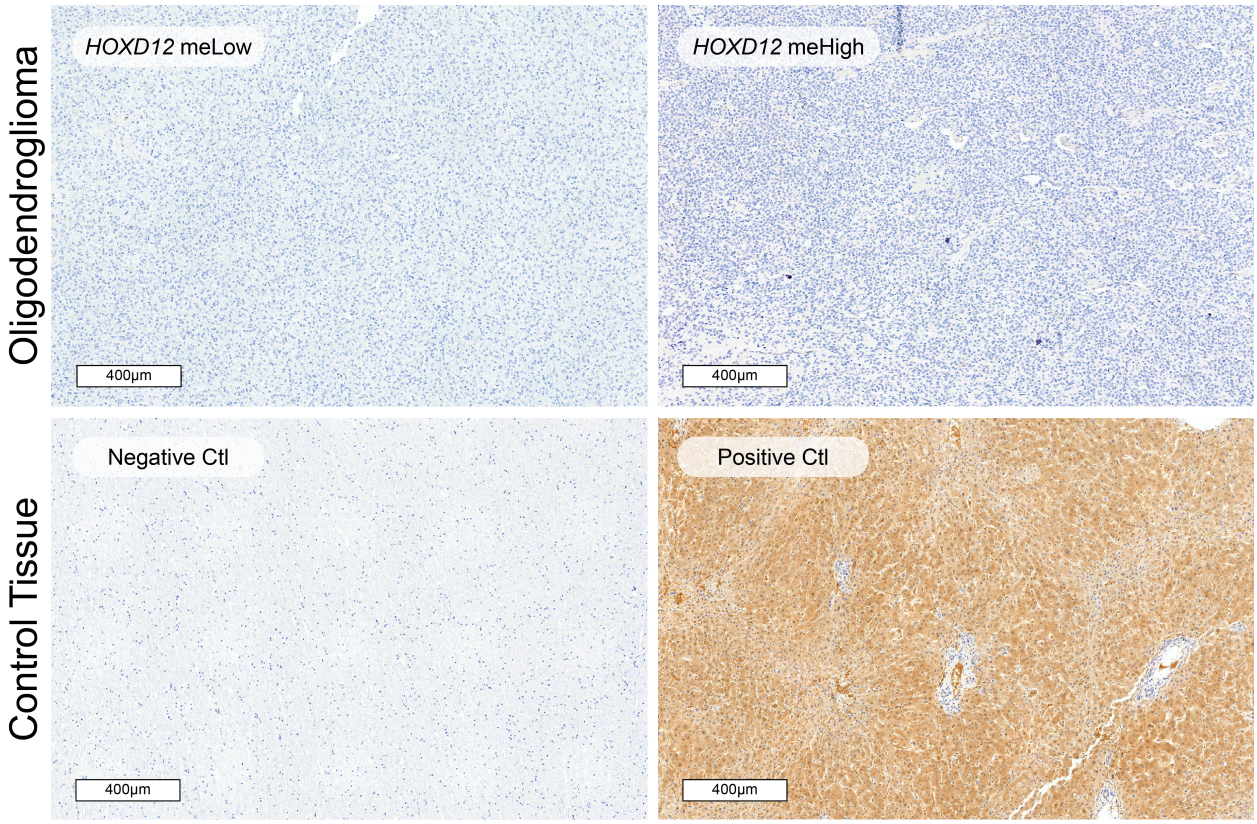


Figure A.24: *HOXD12* immunohistochemistry is not a sensitive surrogate marker for *HOXD12* methylation status. There is lack of detectable *HOXD12* immunostaining in oligodendrogliomas with either relatively hypomethylated (meLow) or hypermethylated (meHigh) *HOXD12*, similar to normal brain (negative control tissue). *HOXD12* is reliably detected in hepatocytes (positive control) with internal negative control (vessels).

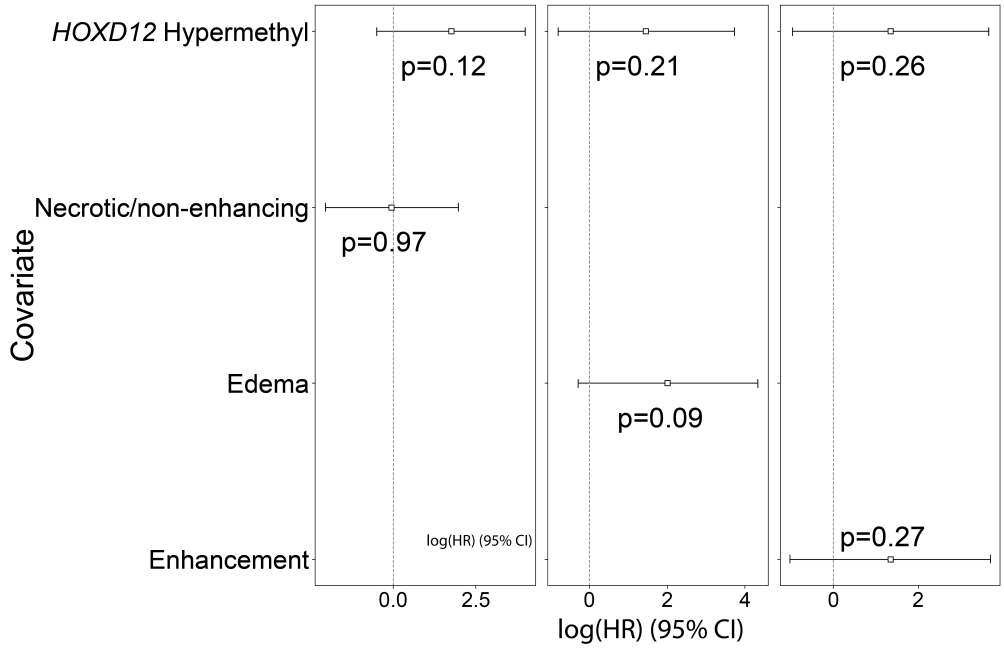


Figure A.25: *HOXD12* hypermethylation was not independently prognostic of radiographic features. Multivariate analyses showed that *HOXD12* hypermethylation was not independently prognostic of tumor necrotic/non-enhancing, edema, or enhancement volume. Whole tumor volume was omitted due to collinearity.

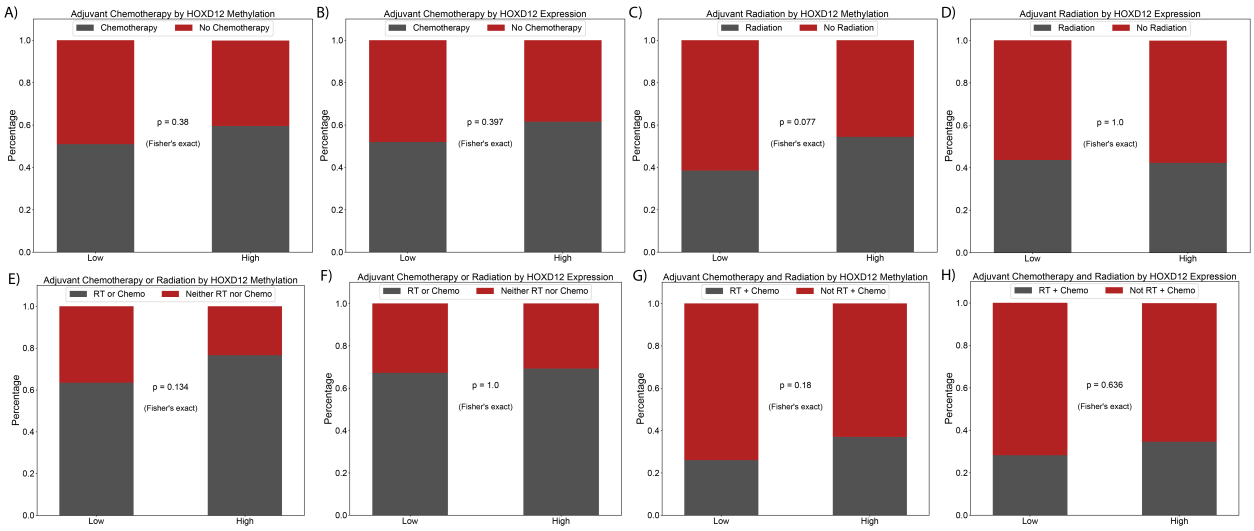


Figure A.26: No significant treatment differences were observed between oligodendroglioma patients with *HOXD12* hypermethylation and/or *HOXD12* overexpression in the TCGA. A) No significant adjuvant chemotherapy treatment difference was observed between patients with or without *HOXD12* hypermethylation. B) No significant adjuvant chemotherapy treatment difference was observed between patients with or without *HOXD12* overexpression. C) No significant radiation treatment difference was observed between patients with or without *HOXD12* hypermethylation. D) No significant radiation treatment difference was observed between patients with or without *HOXD12* overexpression. E) Comparing patients harboring and not harboring *HOXD12* hypermethylation, there was no significant difference in patients who were treated with either adjuvant chemotherapy or radiation therapy vs. neither. F) Comparing patients harboring and not harboring *HOXD12* overexpression, there was no significant difference in patients who were treated with either adjuvant chemotherapy or radiation therapy vs. neither. G) Comparing patients harboring and not harboring *HOXD12* hypermethylation, there was no significant difference in patients who were treated with both adjuvant chemotherapy and radiation therapy vs. those treated with monotherapy or no therapy. H) Comparing patients harboring and not harboring *HOXD12* overexpression, there was no significant differences in patients who were treated with both adjuvant chemotherapy and radiation therapy vs. those treated with monotherapy or no therapy.

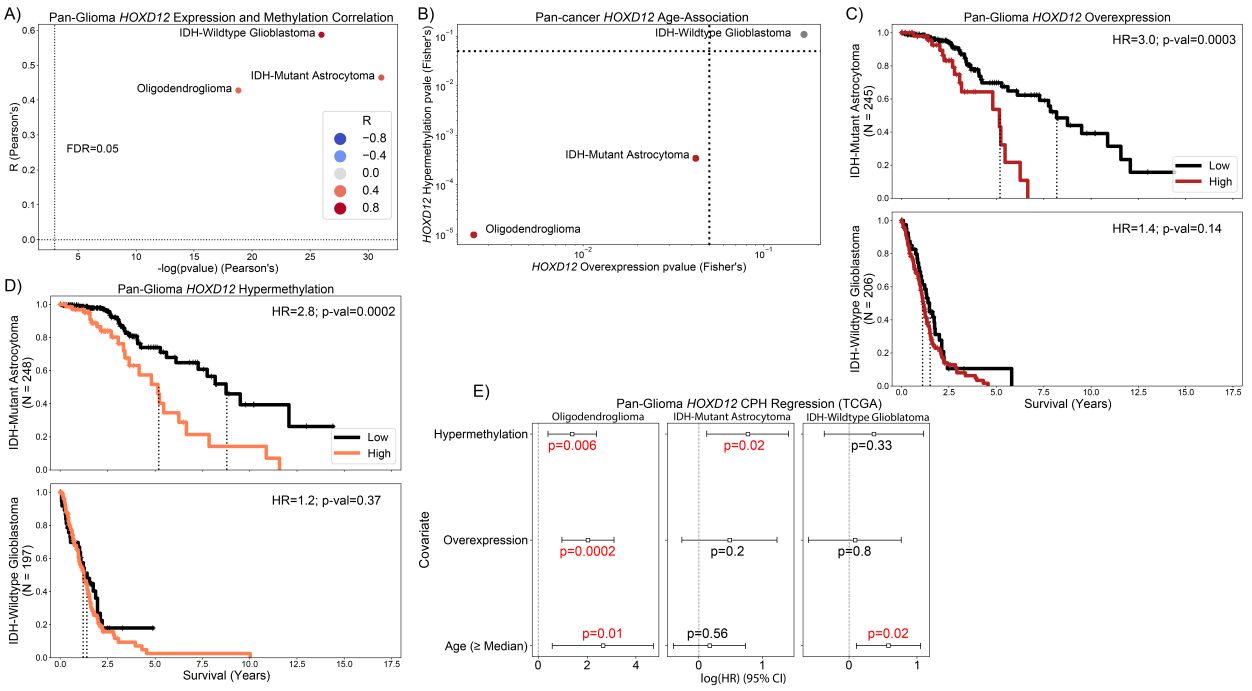


Figure A.27: *HOXD12* expression and methylation levels were correlated in all adult diffuse glioma subtypes but prognostic and age-associated only in IDH-mutant tumors. **A)** *HOXD12* expression and methylation levels were significantly correlated in IDH-mutant astrocytoma and IDH-wildtype glioblastoma in addition to oligodendroglioma. **B)** TCGA IDH-mutant astrocytomas, but not IDH-wildtype glioblastomas, were age-associated. **C)** *HOXD12* overexpression was prognostic in IDH-mutant astrocytoma (p¡0.001, log-rank) but not IDH-wildtype glioblastoma (p=0.14, log-rank). **D)** Similarly, *HOXD12* hypermethylation was negatively prognostic in IDH-mutant astrocytoma (p¡0.001, log-rank), but not IDH-wildtype glioblastoma (p=0.37, log-rank). **E)** As reported in the main text, *HOXD12* overexpression and hypermethylation were both independently prognostic of patient age in oligodendroglioma. *HOXD12* hypermethylation, but not overexpression, was independently prognostic of patient age in IDH-mutant astrocytoma. Neither *HOXD12* overexpression nor hypermethylation were independently prognostic of patient age in IDH-wildtype glioblastoma.

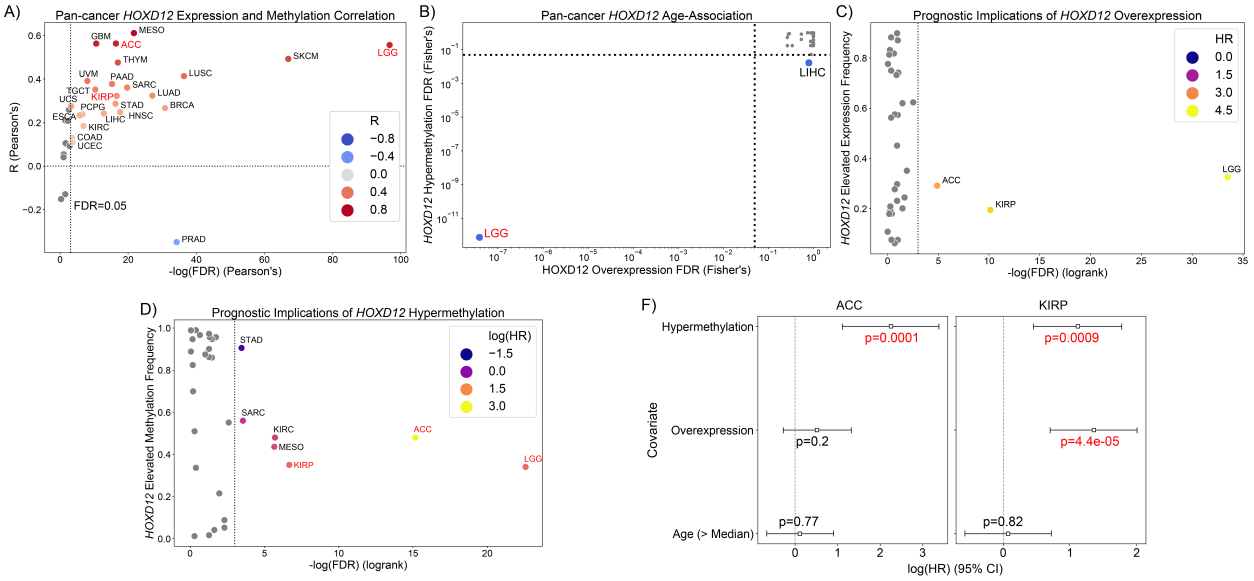


Figure A.28: *HOXD12* expression and methylation levels were correlated, prognostic in selected TCGA cancer types, but not associated with age. **A)** *HOXD12* expression and methylation levels are significantly correlated in 24 of the 33 TCGA cancer types (73%). **B)** The TCGA lower grade glioma (LGG) dataset, which contains almost all TCGA oligodendrogliomas and most TCGA IDH-mutant astrocytomas, was the only TCGA cancer type associated with both *HOXD12* overexpression and hypermethylation. **C)** *HOXD12* overexpression was prognostic in adrenocortical carcinoma (ACC FDR¡ 0.01, log-rank) and renal papillary cell carcinoma (KIRP FDR¡1e-4, log-rank) in addition to low-grade gliomas (LGG FDR¡1e-14, log-rank). **D)** *HOXD12* hypermethylation was also negatively prognostic in ACC (FDR¡1e-6, log-rank), KIRP (FDR¡0.002, log-rank), and LGG (FDR¡1e-9, log-rank) along with kidney renal clear cell carcinoma (KIRC FDR¡0.005, log-rank), mesothelioma (MESO FDR¡0.005, log-rank), and sarcomas (SARC, FDR=0.029, log-rank). **E)** ACC survival was driven primarily by *HOXD12* hypermethylation; *HOXD12* hypermethylation and overexpression were independently prognostic of each other and age in KIRP tumors.

VITA

Nicholas Nuechterlein received his B.S. in Mathematics and English Literature from the University of Michigan in Ann Arbor. He is currently a Ph.D. candidate in Computer Science & Engineering at the University of Washington. His research interests include glioma, bioinformatics, computer vision, and machine learning.