RICE UNIVERSITY

Identifying Image Derived Features of Radiation Therapy Response: Tumor and Normal Tissue

by

Tien T. Tang

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE

Amina Qutub, Chair
Assistant Professor of Bioengineering

M. Waleed Gaber, Director
Associate Professor of Pediatrics-Oncology at BCM

Gang Bao
Foyt Family Professor of Bioengineering

Laura Segatori
Associate Professor of Chemical and Biomolecular Engineering

HOUSTON, TEXAS
December 2018
ABSTRACT

Identifying Image Derived Features of Radiation Therapy Response: Tumor and Normal Tissue

by

Tien T. Tang

Brain tumors constitutes the second most common malignancy in children. Management of these tumors with surgical resection, radiation therapy and chemotherapy presents significant challenges, with cure rates lagging compared to other pediatric cancers. While the introduction of radiation therapy (RT) has significantly improved patient outcome, survivors are never the less prone to cognitive impairment and other radiation-induced side effects. Therefore early detection of treatment resistance and treatment side effects are important for treatment planning and patient prognosis. Monitoring of brain tumor's response is commonly done using medical imaging techniques such as magnetic resonance (MR) and positron emission tomography (PET). In addition to the clinical value of providing information regarding tumor location, size, and metabolism, these images can also be further analyzed to extract quantitative imaging features which can provide additional information for tumor characterization that preserves the spatial and temporal heterogeneity of the tumor.

In this work, texture analysis will be utilized to establish quantitative image features that will assist in understanding and predicting RT response of tumors and
detection of radiation-induced normal tissue injury. Using preclinical models, quantitative image features will be mined from MR and PET scans in radioresponsive and radioresistant tumors to establish universal and tumor-specific imaging markers of treatment response. Furthermore we will establish imaging markers that will provide immediate readout of normal tissue injury and map out the long term changes caused by RT. The outcome of our research will provide clinicians with a toolset to predict, detect, and understand RT response in both tumor and normal tissue for the personalization of treatment for affected children.
Acknowledgments

It takes a village to raise a PhD student:

Dr. Gaber, thank you for your guidance, support, encouragement and allowing me to stay curious.

Dr. Qutub, thank you for your guidance and for allowing me to be part of not only one but two great labs.

Dr. Segatori, and Dr. Bao, thank you for being on my thesis committee.

Janice, thank you for being my Chipolte/Chick-fil-A/McD buddy, for being a great teacher and role model.

Qutub lab, especially Arun(A), Byron, Wendy, Andre and George thank you for making science discussions fun.

None of this would be possible without the support and unconditional love from my family, Mẹ, Cha and Hiền.

And finally David, thank you for being my personal editor, chef, cheerleader and keeping me grounded throughout this journey.
# Contents

Acknowledgments ........................................................................................................ iii

Contents ......................................................................................................................... iv

List of Figures .................................................................................................................. vii

List of Tables ................................................................................................................... ix

List of Equations ............................................................................................................. x

Nomenclature .................................................................................................................. xi

Introduction ..................................................................................................................... 12

Background ..................................................................................................................... 18

  Medulloblastoma ........................................................................................................... 18
  Radioreistance in brain tumors ....................................................................................... 21
  Radiation-induced brain injury ....................................................................................... 24

Medical Imaging .............................................................................................................. 26

  Magnetic resonance imaging ....................................................................................... 27
  Positron emission tomography ..................................................................................... 28

Texture analysis .............................................................................................................. 28

  GLCM image features .................................................................................................. 30

Development of program for feature extraction and calculation .................................... 31

Imaging Predictors of Radioreistance: Sonic Hedgehog-Medulloblastoma model ............ 33

  3.1. Introduction ............................................................................................................. 33

  3.2. Materials and Methods ........................................................................................ 36

    Establishing radioresistant SHH medulloblastoma model ........................................ 36

    In-vitro characterization of DAOY-RR ....................................................................... 36

    Tumor implantation ..................................................................................................... 38

    MR imaging ................................................................................................................ 38

    PET imaging .............................................................................................................. 39

    Clinical data ............................................................................................................... 40

    Image feature extraction ............................................................................................ 40

    Prediction model ........................................................................................................ 41
Statistical analysis ........................................................................................................................................... 41

3.3. Results ..................................................................................................................................................... 42
Radioresistance phenotype established with repeated irradiations .............................................................. 42
Radioresistant phenotype is preserved in-vivo ............................................................................................ 44
Functional changes detected after RT ........................................................................................................ 45
Prediction model performance .................................................................................................................... 48
Clinical data .................................................................................................................................................. 51

3.4. Discussion .............................................................................................................................................. 55

3.5. Conclusions ........................................................................................................................................... 59

Robustness of Radioresistant Image Predictors: Glioma model ................................................................. 60

4.1. Introduction ........................................................................................................................................... 60

4.2. Methods ................................................................................................................................................ 64
Texture feature extraction ............................................................................................................................ 64
Tumor type classification model .................................................................................................................. 65
Tumor growth rate prediction ....................................................................................................................... 66

4.3. Results ................................................................................................................................................... 67
Tumor Type Classification .......................................................................................................................... 69
Growth Curve Prediction ............................................................................................................................. 74

4.4. Discussion ............................................................................................................................................. 76

4.5. Conclusions ......................................................................................................................................... 80

Imaging Markers of Radiation-Induced Brain Injury ................................................................................. 81

5.1. Introduction ......................................................................................................................................... 81

5.1. Methods ............................................................................................................................................... 84
Establishing radiation-induced brain injury model: late RT effects .............................................................. 84
Cognitive Testing ........................................................................................................................................ 85
Imaging – Diffusion tensor imaging ............................................................................................................ 86
Imaging – T_1-weighted/T_2-weighted/Proton Density scans ........................................................................ 87
Imaging – Positron emission tomography ................................................................................................... 88
Immunohistochemistry ............................................................................................................................... 89
Statistical analysis ...................................................................................................................................... 89

5.2. Results .................................................................................................................................................. 90
Brain growth .......................................................................................................................... 90
Cognitive performance ......................................................................................................... 93
Diffusion tensor imaging data .............................................................................................. 96
White matter .......................................................................................................................... 98
Imaging feature and cognitive performance correlation ......................................................... 99
Inflammation .......................................................................................................................... 101
5.3. Discussion ...................................................................................................................... 104
5.4. Conclusions .................................................................................................................... 108
Conclusion ............................................................................................................................ 109
Appendix A ............................................................................................................................ 112
Appendix B ............................................................................................................................ 115
Appendix C ............................................................................................................................ 118
References ............................................................................................................................. 122
List of Figures

Figure 1. Skematic of the overall thesis work. ................................................................. 16
Figure 2. Summary of medulloblastoma subgroups\textsuperscript{16} ........................................ 20
Figure 3. Skematic of GLCM construction from a gray scale image. ................................. 30
Figure 4. GUI developed for tumor segmentation and texture feature extraction ............ 32
Figure 5. In-vitro characterization of radioresiant DAOY (DAOY-RR) and wild-type DAOY
(DAOY-WT) cell lines ........................................................................................................ 43
Figure 6. Tumor growth curve............................................................................................. 45
Figure 7. Functional imaging of DAOY-WT and DAOY-RR pre and post-RT .................. 47
Figure 8. Validation accuracy of each classification model ................................................ 49
Figure 9. AUC of classification model ................................................................................ 50
Figure 10. Example of the images used for feature extraction ........................................... 53
Figure 11. Measures of classifier performance. ................................................................. 54
Figure 12. T1-weighted post contrast MR scan of A. GL261, mouse glioma, B. U87, human
glioma .................................................................................................................................. 63
Figure 13. T\textsubscript{1} and T\textsubscript{2} relaxation time of glioma and medulloblastoma tumors........... 69
Figure 14. Validation accuracy of each classification model of given gray level .................. 70
Figure 15. Importance of features in tumor classification .................................................... 73
Figure 16. Tumor growth prediction model ....................................................................... 75
Figure 17. Functional imaging changes at 48 hours post-RT ............................................. 83
Figure 18. Experimental time line for radiation-induced brain injury model ....................... 85
Figure 19. RT stunts brain growth. ....................................................................................... 91
Figure 20. Volumetric data from DTI scans demonstrate variation radiation response in
different tumor regions ................................................................................................. 92
Figure 21. 5-CSRTT testing results at 12 months post-RT..........................................................94

Figure 22. 5-CSRTT testing results when tested at Stage 7. ....................................................95

Figure 23. DTI reveals changes in brain regions and fiber values after RT. ..............................97

Figure 24. RT affects global connectomes. Error bars indicate 95% CI. Connectome metrics were calculated from DTI imaging at 12 months post-RT. .........................................................98

Figure 25. White matter analysis with standard MR images. ..................................................99

Figure 26. Correlation matrix of image features and 5-CSRTT performance. .......................100

Figure 27. At 12 months post-RT there was significant increase in metabolic activity in the brain.................................................................................................................................102

Figure 28. Neuroinflammation and demyelination is observed up to 12 months post-RT......103

Figure 29. Presence of microglia is significantly higher in RT brains....................................104
List of Tables

Table 1. List of molecular markers of tumor radioresistance found to be associated with radioresistance. ........................................................................................................................................ 23

Table 2. Parameters derived from the linear-quadratic model characterizing cell’s response to RT, ****p≤0.0001 ........................................................................................................................................ 42

Table 3. Metabolic activity of tumor region as measured by 18F-FDG PET imaging. ................. 46

Table 4. Features that differentiated tumor types, GLCM with 512 gray levels. ....................... 68

Table 5. Classification model performance for GL261 in different tumor regions................. 71

Table 6. Classification model performance for U87 in different tumor regions. ..................... 72

Table 7. Classification model performance for Daoy in different tumor regions. ................... 72

Table 8. Mean SUV of different brain regions. Data is shown as mean ± standard deviation. 103
List of Equations

(1) $P_{i,j}$; $\delta, \alpha$......................................................................................................................29

(2) $PE = \text{No. of colonies formed No. of cells seeded} \times 100\%$........................................37

(3) $SF = \text{No. of colonies formed after treatment No. of cells seeded} \times PE$ .............37

(4) $SF = e - \alpha \text{ dose} - \beta \text{ dose}^2$.........................................................................................37

(5) tumor volume $= ae\beta \times \text{time}$.................................................................................................66

(6) Accuracy $= TP + TNTP + FP + FN + TN$ ..................................................................................66

(7) Specificity $= TNFP + TN$ ..............................................................................................................66

(8) Sensitivity $= TPTP + FN$ ...............................................................................................................67

(9) $F\text{score} = 2 \times TP^2 \times TP + FP + FN$ .....................................................................................67
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-CSRTT</td>
<td>5-Choice Serial Reaction Time Task</td>
</tr>
<tr>
<td>ADC</td>
<td>apparent diffusion coefficient</td>
</tr>
<tr>
<td>ASL</td>
<td>arterial spin labeling</td>
</tr>
<tr>
<td>AUC</td>
<td>area under the curve</td>
</tr>
<tr>
<td>CSC</td>
<td>cancer stem cells</td>
</tr>
<tr>
<td>CSF</td>
<td>cerebrospinal fluid</td>
</tr>
<tr>
<td>CT</td>
<td>computed tomography</td>
</tr>
<tr>
<td>DICOM</td>
<td>digital imaging and communications in medicine</td>
</tr>
<tr>
<td>DTI</td>
<td>diffusion tensor imaging</td>
</tr>
<tr>
<td>FA</td>
<td>Fractional anisotropy</td>
</tr>
<tr>
<td>FBS</td>
<td>fetal bovine serum</td>
</tr>
<tr>
<td>FDG</td>
<td>fluorodeoxyglucose</td>
</tr>
<tr>
<td>FLAIR</td>
<td>fluid-attenuated inversion recovery</td>
</tr>
<tr>
<td>GLCM</td>
<td>gray level co-occurrence matrix</td>
</tr>
<tr>
<td>GUI</td>
<td>graphical user interface</td>
</tr>
<tr>
<td>IACUC</td>
<td>Institutional Animal Care and Use Committees</td>
</tr>
<tr>
<td>IQ</td>
<td>intelligence quotient</td>
</tr>
<tr>
<td>LCA</td>
<td>large cell and anaplastic</td>
</tr>
<tr>
<td>MBEN</td>
<td>medulloblastoma with extensive nodularity</td>
</tr>
<tr>
<td>MRI</td>
<td>magnetic resonance imaging</td>
</tr>
<tr>
<td>msec</td>
<td>millisecond</td>
</tr>
<tr>
<td>PET</td>
<td>Positron emission tomography</td>
</tr>
<tr>
<td>ROI</td>
<td>region of interest</td>
</tr>
<tr>
<td>RR</td>
<td>radioresistant</td>
</tr>
<tr>
<td>RT</td>
<td>radiation therapy</td>
</tr>
<tr>
<td>SHH</td>
<td>sonic hedgehog</td>
</tr>
<tr>
<td>SUV</td>
<td>standardized uptake value</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
<tr>
<td>WT</td>
<td>wild-type</td>
</tr>
</tbody>
</table>
Introduction

Brain tumors constitute the second most common malignancy in children with medulloblastoma, accounting for 20% of these tumors\textsuperscript{1,2}. Despite an aggressive treatment regimen which includes surgical resection, radiation therapy (RT) and chemotherapy, the cure rates for brain tumors lag compared to the successes achieved in treating other pediatric malignancies\textsuperscript{3,4}. Furthermore, pediatric patients are prone to acute- and late-radiation-induced brain injury which can lead to long term cognitive impairment, affecting memory, attention and learning ability\textsuperscript{5,6}. Cancer diagnosis and treatment planning often rely on the molecular information such as sonic hedgehog (SHH) expression and histological features, obtained from the resected tumor tissue. However, tumor biopsies are limited in that they only provide information about a small subset of the tumor at a single snapshot in time, and thus cannot fully capture the spatial and temporal intratumoral heterogeneity. Furthermore not all tumors are resectable. Therefore an alternative technique that
can noninvasively establish diagnoses and prognoses, and can further be used to monitor treatment response, would greatly improve clinical outcomes for brain tumor patients.

Medical imaging techniques commonly used for the diagnosis and detection of cancer such as magnetic resonance (MR) and positron emission tomography (PET) are readily available in the clinic. These imaging modalities are primarily used for anatomical purposes such as tumor localization, tumor volume and detection of metastatic disease. However, recent studies have shown that in addition to anatomical information, phenotypic assessments, such as midline shift or degree of tumor border enhancement, are reflective of the tumor’s molecular activity.

Although these phenotypic assessments are performed by trained radiologists, they can be highly subjective and not easily reproducible. The reproducibility of these assessments can be improved using techniques from the emerging field of radiomics, which aims to extract quantitative image features in a high throughput manner from phenotypic patterns in diagnostic images, particularly including features that are not visually evident or easily quantified such as cluster prominence, variation in gray-scale, or heterogeneity. Using this approach, texture analysis can be performed to extract new imaging features from MR and PET images that can provide complementary or even interchangeable information with other clinical data, which can be easily integrated into the current clinical workflow.
Identification of these imaging markers may help identify radioresponsive and radioresistant tumor phenotypes, which will in turn provide clinicians with an additional toolset to detect, predict, and monitor treatment response in order to personalize the treatment of affected children. In this regard, the hypothesis of the proposed work is that distinct imaging features can be extracted from medical imaging techniques that can be used to predict radiation therapy outcome and evaluate radiation-induced normal tissue injury. This hypothesis will be tested in the following specific aims:

**Aim 1: Establish imaging signatures for radiation responsive and resistant tumors in a Sonic Hedgehog-Medulloblastoma model. In this aim, we will:**

A. Establish a radioresistant Sonic Hedgehog-Medulloblastoma model *in vivo*

B. Identify imaging signatures of a radioresistant and radioresponsive tumor in PET and MR images

C. Construct a prediction model using these signature to identify radioresistant tumor response early in treatment

D. Validate the prediction model of treatment response using clinical MR images

**Aim 2: Determine the robustness of radioresistant imaging signatures in glioma tumors. In this aim, we will:**

A. Determine the applicability of the established radioresistance signatures in a different cell line – a glioma tumor model

B. Identify diagnostic imaging signatures in glioma and medulloblastoma models
Aim 3: Establish imaging signatures of radiation-induced brain injury after cranial irradiation. In this aim, we will:

A. Determine imaging markers of acute and late term radiation-induced tissue damage in PET and MR images

B. Identify image features that are indicative of cognitive impairment from long term brain injury

Upon completion of this proposal a workflow of region of interest segmentation, feature extraction and construction of predictive models will be established that can be utilized for understanding treatment response and normal tissue injury following radiation therapy (Figure 1). This work further sheds light onto potential novel imaging biomarkers which may impact the clinical workflow as well as basic discovery research.
Figure 1. Schematic of the overall thesis work. Image features extracted from PET/CT and MR are used to construct a feature matrix where each row represents a response and its corresponding feature value. In **Aim 1** the feature matrix is used to construct a prediction model of treatment outcome. In **Aim 2** the matrix is used to identify imaging markers of medulloblastoma (MB) and glioma tumors. In **Aim 3** the feature matrix and functional imaging is used to detect imaging markers of radiation-induced brain injury and their correlation to cognitive performance.
The chapters are organized as follows:

Chapter 1 is an overview of the thesis and the specific aims that will be addressed.

Chapter 2 provides background information on medulloblastoma, clinical significance of radiation therapy treatment and the developed software for texture analysis of medical images.

Chapter 3 presents a new radioresistance preclinical medulloblastoma model and the use of the model to identify imaging markers and predictive model of treatment resistance. Clinical utility of prediction model is established with medical images acquired at Texas Children's Hospital.

Chapter 4 demonstrates the utility of image features to diagnosis brain tumors and predict growth rate with presurgical images.

Chapter 5 presents new imaging markers of late term radiation-induced brain injury and the significance of these imaging markers on cognitive performance.

Chapter 6 provides a discussion of the work as a whole and the impact these findings could have on future clinical practices.
Chapter 2

Background

Radiation therapy (RT) is a vital part in the current treatment protocol for brain cancer. However, RT efficacy is limited by treatment resistance and by tolerance of normal tissue to ionizing radiation. For pediatric patients, especially those under the age of three, due to the developmental consequences RT can have on the developing brain, a better understanding of RT response is needed to help clinicians tailor the role of radiation in treatment planning\textsuperscript{14}. This chapter will provide an overview of the role that radiation therapy has in the current standard of care for medulloblastoma (one of the most common pediatric brain tumor), challenges of RT treatment resistance or radioresistance, and the role different medical imaging techniques may have in predicting treatment response.

Medulloblastoma

Medulloblastoma is an invasive embryonal tumor of the cerebellum occurring predominantly in the pediatric population. According to the World Health Organization (WHO), medulloblastoma is classified into pathological and genetic
subgroups. Classical pathological classifications include: classic medulloblastoma, desmoplastic/nodular, medulloblastoma with extensive nodularity (MBEN), large cell, and anaplastic medulloblastoma. These groups can be further combined with overlapping morphologies into large cell and anaplastic (LCA). With advances in high throughput genomic screening these tumors have been further classified according to their molecular subtype which includes: Wnt, which is identified by nuclear accumulation of β-catenin through upregulation of canonical Wnt signaling; sonic hedgehog (SHH) which is identified by activation of the SHH-signaling pathway; Group 3; and Group 4. Figure 2 is a summary as presented by Taylor et al of the pathological and histological characteristic of each medulloblastoma subgroup. While both Wnt and SHH are associated with different germline mutations, the underlying biology of the other two subgroups, Group 3 and Group 4, is not well understood although both groups contain widespread chromosomal and transcriptional alterations that are clustered into these two separate subgroups. Due to this lack of definitive profile, new diagnosed medulloblastomas in Group 3 and Group 4 are currently classified based only on clustering to the previously established subgroups. However, recent studies with whole genome sequencing have identified potential subgroup specific driver mutations including Notch and TGFβ signaling which may provide new metric for classifying these subgroups.
Identification and classification of these medulloblastoma subgroups have important clinical implications since the prognosis differs drastically across these groups. Group 4, the most common (accounting for 35% of all diagnosed tumors), and the SHH subgroup both have a 5 year survival rate of approximately 75% with the current standard of care that includes: surgical resection, risk-adaptive therapy and adjuvant chemotherapy. The clinical significance of these subgroups is most clearly demonstrated by the differences in survival rates of the Wnt subgroup and Group 3, where 5 year survival is 90% for the Wnt subgroup but only 50% for
Group 3. Furthermore, 20% of children classified in the SHH subgroup present with metastatic disease at the time of medulloblastoma diagnosis compared to 45% in Group 3, which further complicates prognostic determination and treatment planning. The patient age demographics also vary with subgroup, for instance the SHH subgroup forms a bimodal distribution with age, with high incidences in infants and older children, unlike the other subgroups where prevalence is high in younger children. Even with such drastic differences between these groups, in terms of prognosis and treatment outcome, treatment plans are still based only on the patient age at the time of diagnosis and on the presence of metastatic disease.

To further complicate treatment planning, the biological processes responsible for the heterogeneous response to therapy within the same molecular subgroup are still unknown, and thus there is currently no subgroup-specific standard of care. Although ongoing clinical trials aim to incorporate this molecular classification into treatment planning, it is NOT yet been integrated into the standard of care.

**Radioresistance in brain tumors**

The introduction of RT in the mid 60's has greatly increased the survival rate of pediatric brain tumor patient from 54% to 71% compared to surgical resection alone. RT uses high-energy radiation with x-ray, gamma rays or charged particles to kill cancer cells through DNA damage. Due to cancer cells proliferating at a higher rate than normal cells, they are generally more vulnerable to DNA damage, which allows for effective local tumor control while minimizing harm to the surrounding...
normal tissue\textsuperscript{27}. With current treatment plans standard risk and high risk patients, as delineated by the absence or presence of metastatic disease, are given a radiation dose of 23-36 Gy to the craniospinal axis with a boost dose of radiation to the primary tumor site that accumulates to 54-56 Gy. This is then followed by multiple rounds of chemotherapy in varying combination of vincristine, cisplatin, cyclophosphamide\textsuperscript{23}. Despite the aggressive treatment regimen, not all patients will respond to treatment.

Radioresistance, defined as the lack of response with minimal cell kill after exposure to ionizing radiation, manifests clinically as progressive disease, partial response, or disease recurrence\textsuperscript{28}. This phenomenon can be caused by intrinsic factors, cellular or microenvironmental factors, or treatment-induced selective processes\textsuperscript{29,30}. Although different genes have been associated with radioresistance in many tumor types (Table 1), the exact mechanisms that drive medulloblastoma and tumors in general to radioresistance are still a subject of intensive research.

The leading approach to identify biological processes responsible for tumor response has been the \textit{bottom-up} approach that aims to correlate genetic and molecular profiles from patient biopsies to treatment outcomes from clinical data\textsuperscript{24,31-35}. From these studies cancer stem cells (CSCs) have become a potential candidate for tumor’s resistance to treatments that target DNA damage pathways such as RT\textsuperscript{36-39}. CSCs found in glioma, a highly radioresistant brain tumor, have been shown to have enhanced DNA repair capacity and cell cycle arrest\textsuperscript{37,40}. An alternative proposed mechanism for the emergence of CSCs is the repopulation and
redistribution of these cells after RT. Ionizing radiation can lead to a switch to a symmetric division in which the stem cells produces two daughter stem cells rather than one daughter cell and one differentiated cell\textsuperscript{41}. The expansion of the CSCs subpopulation and their insensitivity to RT can cause the tumor to shift towards a radioresistant phenotype in a spatially heterogeneous manner. CSCs also strive in perivascular niches, which implicate the microenvironment in treatment resistance, further demonstrating the complex relationship between intrinsic cell population and the tumor’s microenvironment\textsuperscript{36,42}.

**Table 1.** List of molecular markers of tumor radioresistance found to be associated with radioresistance.

<table>
<thead>
<tr>
<th>Radioresistance genes</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>SFP1, SFP2, FGF-2, ATM, NBS1, MRN complex, ATM, ATR, Chk1, Chk2, pleiotropic β (TGF-β), RhoB, IGF1, MAPK, Bcl-2, Bcl-XL, ASCL1, FRAT1, STAT3, MTOR, MET, EGFR, Notch ligands (DLL1, DLL3)</td>
<td>Kelley et al 2016\textsuperscript{43}</td>
</tr>
<tr>
<td>P2X7</td>
<td>Gehring et al 2015\textsuperscript{44}</td>
</tr>
<tr>
<td>SAT1</td>
<td>Brett-Morris et al 2014\textsuperscript{45}</td>
</tr>
<tr>
<td>Tp53, wtp53</td>
<td>D’Avenia et al 2006\textsuperscript{46}</td>
</tr>
<tr>
<td>HIF1</td>
<td>Moeller et al. 2004\textsuperscript{47}</td>
</tr>
<tr>
<td>TopBP1, CLSPN</td>
<td>Choi et al 2014\textsuperscript{48}</td>
</tr>
</tbody>
</table>

Consequently, while the preferred workflow for prediction of treatment response might involve genetic screening and the identification of genes that drive treatment response, this is far from being integrated into current clinical practices. Furthermore, generating these genetic profiles would rely heavily on tumor biopsies which lack both temporal and spatial resolutions and are only available if the tumor
is accessible. Combined with tumor heterogeneity, these complications render the bottom-up approach insufficient to understand the dynamic molecular network that drives RT response and the identification of the resistance phenotypes. Faster, less expensive and more importantly noninvasive techniques are desirable in identifying treatment resistance and develop treatment plans accordingly.

To this end, many types of medical images already acquired as part of the standard clinical workflow are readily available for further analysis towards predicting treatment response. Identification of radioresistance or radiosensitive imaging markers in standard diagnostic medical images, such as MR or PET scans, can allow for an easily implemented clinical workflow in which clinicians can make more informed treatment decisions based upon the expected treatment response. This will be especially beneficial for pediatric patients where RT can have long term side effects from radiation-induced brain injury.

**Radiation-induced brain injury**

While RT has greatly improved the outcome of childhood brain cancer, long-term survivors often suffer from radiation-induced brain injury. In particular, survivors of childhood brain cancer treated with RT have a significantly lower intelligence quotient (IQ) compared to survivors with only surgical resection\(^6\). Children treated for medulloblastoma using radiation experience a mean 1.0 standard deviation (15 IQ points) decrease during a 4-year period\(^6\), which corresponds to a decline from the 50\(^{th}\) to the 16\(^{th}\) percentile in global cognitive functioning. This substantial, irreversible change has major, life-long implications
for educational achievement and quality of life. Radiation-induced brain injury can also lead to hormonal deficiencies, growth decline as well as secondary malignances. Even though radiation treatment plans optimize dose deposition to the primary tumor site while sparing normal tissue, it is inevitable that the normal tissue will receive some radiation dose and this has been shown to result in neurocognitive deficit in a dose-dependent manner.

Multiple mechanisms have been implicated in causing radiation-induced brain injury. White matter degradation is considered to be a primary marker of injury, as pediatric brain tumor survivors treated with RT have found to have significantly less white matter volume. Neuroinflammation, both acute and chronic, have also been implicated in driving the cognitive outcome of these survivors where initial damage from RT can prompt acute activation of microglia, which in turn may persist for months or years after RT due to secretion of inflammatory cytokines and the production of reactive oxygen species. Preclinical studies have found that post-RT inflammation of the microvascular environment can lead to impaired memory up to 9 months post-treatment. While neuroinflammation may be a common factor across radiation induced brain injuries, other molecular mechanisms, including increased apoptosis of endothelial cells leading to break down of the blood brain barrier, diminishing neurogenesis in the hippocampus, and alteration of neuronal synaptic morphology, function and connections are also highly relevant.
Currently, there are no established molecular markers to predict cognitive outcome of pediatric brain tumor survivors although age at time of RT has been found to correlate strongly with cognitive decline – namely that younger patients experience greater cognitive decline. It is thought that at a younger age, the brain is still undergoing rapid growth and development and is therefore more sensitive to radiation damage. Accordingly, for children diagnosed under the age of 3, chemotherapy is often used to delay RT until the child is older, or eliminate the use of radiotherapy altogether. However, even when RT is successfully delayed using varying doses of cyclophosphamide, cisplatin, vincristine sulfate and etoposide phosphate, patients will nevertheless have permanent cognitive deficits.

Even though older patients do not show the same extent of overall cognitive decline they still face learning disabilities ranging from deficits in learning, executive function, attention and memory. While these cognitive impairments are commonly observed there is currently no universal model of the molecular mechanisms that drives this cognitive decline and it is more unclear what drives heterogeneity of these responses. The identification of molecular markers or prognostics marker of this cognitive decline can help the clinicians make more informed treatment plans and provide additional care to mitigate these side effects.

**Medical Imaging**

Medical imaging is one approach to overcome some of the limitations of using biopsy samples to understand treatment response. Currently in the clinic,
imaging is primarily used to assess tumor volume and location to determine feasibility of surgical resection. Furthermore, patients are stratified into the two risk groups (standard and high risk) based on those with disseminated or residual disease and those without\textsuperscript{57}. Imaging modalities such as positron emission tomography (PET) and magnetic resonance imaging (MRI) can further characterize medulloblastomas and assess response.

**Magnetic resonance imaging**

Magnetic resonance imaging (MRI) is commonly used in the clinic and is a noninvasive technique that visualizes soft tissue. MRI is based on the reorientation of the nuclear spins of hydrogen atoms in response to magnetic pulses. This is currently part of the clinical workflow that determines the feasibility of surgical resection, risk stratification and to assess the response to treatment based on reduction of tumor volume and/or lack of recurrence. Tumors can easily be delineated with a T\textsubscript{1}-weighted post contrast scan and have a heterogeneous appearance on both T\textsubscript{1}-weighted and T\textsubscript{2}-weighted MR scans. On a T\textsubscript{2}-weighted scan, the tumor will usually have intermediate intensity between white and gray matter which is one of the features radiologist look for to differentiate medulloblastoma from other brain tumors\textsuperscript{58}. Beyond the standard use of MR images in identifying and delineating tumors, current advancements in molecular and functional MR imaging allow for radiological assessments of perfusion rate, metabolic activity and proliferative activity which provide additional information concerning tumor development and response. These advancements highlight the potential for MR
images to yield information pertaining to treatment outcome, beyond simple identification of the tumor.

**Positron emission tomography**

Nuclear medicine-based imaging techniques such as PET imaging with fluorodeoxyglucose ($^{18}$F-FDG) – which is often integrated with computed tomography (CT) imaging – is frequently used for the detection of various cancers based upon their metabolic activity, and later used to monitor the treatment response$^{59-61}$. Functional information such as glucose activity, hypoxic state and proliferative activity acquired using PET imaging can also be indicative of patient outcome$^{62,63}$. Such images could be obtained prior to surgical resection and would serve to inform clinicians about the tumor biology and the anticipated clinical course.

Recent studies have investigated the applicability of medical imaging to predict prognosis and classify the different medulloblastoma subgroups with some success$^{8,64-66}$. Together, these imaging modalities can characterize the tumor metabolism without losing spatial and temporal information.

**Texture analysis**

In order to extract these additional information from medical images, we must employ a novel approach to analyze the images and extract information beyond what is readily apparent to the clinician when they diagnose the tumor.
One approach to extracting quantitative image features from medical images is using texture analysis. Texture analysis is a way to characterize intuitive qualities observed visually such as smoothness, coarseness and uniformity based on spatial organization of pixel intensities\textsuperscript{12,13}. This can be done by constructing a gray level co-occurrence matrix (GLCM) and subsequent extraction of statistical measures from the matrix to describe the image’s texture. The GLCM at its basic level is a histogram, where the frequency in which pairs of pixels appear together is tabulated. The matrix provides insight into the inter-pixel relationship and spatial dependences of the gray levels. The GLCM is defined, for a given N, as:

\begin{equation}
P(i,j; \delta, \alpha)
\end{equation}

Where N is the matrix size and defined by the number of discrete gray level intensities, each (i,j) element represent the frequency in which the combination of intensity level i and j occur as separated by pixel distance (δ) in direction (α). For a 2-dimensional image α=0\(^{\circ}\), 45\(^{\circ}\), 90\(^{\circ}\), and 135\(^{\circ}\) which accounts for symmetry and can also be extended to 3-dimensional analysis. \textbf{Figure 3} is an example of how the GLCM is constructed from a 4 gray levels image where \(\delta=1\) and \(\alpha=0^{\circ}\) with symmetry.

By extracting different statistical measures from the matrix, the texture of the image can be analyzed in a quantitative manner. When coupled with cellular and molecular data, these medical images have the potential to capture important features noninvasively which are informative of treatment response and reflective of the underlying biological processes of the tumor.
Figure 3. Schematic of GLCM construction from a gray scale image. Pixel values 1 and 2 appear adjacent to each other twice in the left image. Correspondingly, element 2,1 of the GLCM is populated with a 2. Note that element 1,2 of the GLCM is also 2 by symmetry, shown by blue arrow.

**GLCM image features**

The statistical parameters or texture features derived from the GLCM describe the spatial relationship between pairs of pixels, and are sensitive to direction ($\alpha$) and distance away from the pixel of interest ($\delta$). The full list of texture features are listed in Appendix A. To provide further clarity in the relationship between what these features represent, listed below are the features and a brief description of what parameter it aims to described in the image$^{67}$.

**Contrast**: measure of local variation in an image, where it is a direct measurement of intensity contrast between the pixels of interest and their neighbors.
**Correlation:** measure of the linear dependency of gray levels of the image, or how correlated a pixel's intensity is to its surrounding neighbors

**Homogeneity:** measures the similarity of pixel intensities

**Energy:** measures uniformity of an image’s pixel intensities

**Entropy:** measures of randomness of an image’s pixel intensities

**Development of program for feature extraction and calculation**

To extract these image features from MR and PET images in a high throughput manner, a software program with a user friendly graphical user interface (GUI) was developed in MATLAB (Figure 4). The program allows the user to either import a single image slice or perform batch analysis on multiple scans. The program can read different image types, but our work focused on using the most commonly used digital imaging and communications in medicine (DICOM) format. Once files are imported by the software, the user can manually segment the region of interest (ROI) from the displayed image. The selected ROI is then used to calculate the different image features with the specified N, α and δ values. The calculated texture features are saved into Excel files as well as into variables containing the segmented ROI, mask and the features associated with each ROI.
Figure 4. GUI developed for tumor segmentation and texture feature extraction. The imported image appears on the left panel where user can manually segment the region of interest. Image processing options appearing on the right side of the GUI.
3.1. Introduction

Medulloblastomas is the most common type of brain tumor\textsuperscript{1,2}. The standard of care for children diagnosed with medulloblastomas consists of chemotherapy, surgical resection where possible, and the mainstay of treatment, radiation therapy (RT). The treatment is prescribed based on a few clinical factors including: presence or absence of metastatic disease, post-operative residual disease and the patient’s age. With current treatment protocols, the overall cure rates are a modest 70-75\%\textsuperscript{3,4}, but nevertheless, approximately 150 children die each year due to treatment failure\textsuperscript{1,5,7}. 

Medulloblastomas have four known distinct molecular subgroups: Wnt, SHH, Group 3 and Group 4\textsuperscript{16-19}. The clinical significance of these subgroups is clearly demonstrated by the differences in survival rates of each subgroup at 5 years, reaching a high of 90\% for the Wnt subgroup but only a meager 50\% for Group 3\textsuperscript{2,21,22}. Furthermore, 20\% of children classified in the SHH subgroup present with metastatic disease at the time of diagnosis compared to 45\% in the Group 3\textsuperscript{22}. The biological processes responsible for the heterogeneous response to therapy within the same molecular subgroup are still unknown, and there is currently no subgroup-specific standard of care\textsuperscript{24,25}.

The current treatment protocol of RT and chemotherapy can lead to long-term side effects such as impaired cognition and stunted growth, however not all tumors will respond to the RT treatment\textsuperscript{56}. Radioresistance, or resistance to RT, manifests clinically as progressive disease, partial response, or disease recurrence\textsuperscript{28}. The exact mechanisms that drive medulloblastoma and tumors in general to radioresistance are still unknown. However, if nonresponsive tumors could be identified before the radiation is delivered, it would be possible to spare these children of the devastating side effects of the treatment and allow alternative treatments options to be explored earlier. By current standards, treatment response however can only be determined after completion of the treatment in which the patient is monitored for relapse\textsuperscript{3,68}. Therefore faster, less expensive and more importantly, noninvasive techniques are desired to establish more accurate prognosis in these pediatric patients to prescribe personalized treatment plans.
Currently in the clinic, imaging is primarily used to assess tumor volume and location, to determine the feasibility of surgical resection, and to assess the response to treatment based on the reduction of tumor volume and/or lack of recurrence. With advancements in molecular and functional imaging, radiological assessments are providing additional information concerning tumor development and response, such as perfusion rate, metabolic and proliferative activities. Such images could be obtained prior to surgical resection and would serve to inform practitioners about the tumor biology and the anticipated clinical course. Recent studies investigated the applicability of medical imaging to predict prognosis and have classified the different medulloblastoma subgroups with success\textsuperscript{8,64-66}. Several studies have attempted to bridge the gap between imaging and molecular activity by correlating imaging features to cellular and gene activities \textsuperscript{8,9,11}.

One approach of quantifying imaging features is texture analysis. Texture analysis is a way to characterize intuitive qualities observed visually based on spatial organization of pixel intensities\textsuperscript{12,13}. This can be done by constructing a gray level co-occurrence matrix (GLCM), where the frequency in which pairs of pixels appear together in a particular offset are considered. By extracting different statistical measures from the matrix, the texture of the image can be analyzed in a quantitative manner. These features, while not intuitively visible, can expand the use of readily available clinical images to predict and understand treatment outcome.
In this study we aim to identify imaging markers of RT response using standard medical images with texture feature analysis. A radioresistant medulloblastoma cell line (DAOY-RR) was established and characterized both in-vitro and in-vivo. Imaging features derived from standard T1-weighted post-contrast scan were used to construct prediction models of RT response. Furthermore the feasibility of using standard medical images in the clinic was verified.

3.2. Materials and Methods

Establishing radioresistant SHH medulloblastoma model

Daoy cells (provided by Dr. Stephen Gottschalk, cells were transfected with a luciferase reporter and GFP) and cultured in Dulbecco’s Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum (FBS) (Sigma-Aldrich, St. Louis, MO) and 1% penicillin streptomycin (Sigma-Aldrich, St. Louis, MO). Cells were grown to ~50% confluence and the Daoy-radioresistant (DAOY-RR) cell line were irradiated in 2 Gy fractions to a total cumulative dose of 50 Gy (RadSource 2000, 25mA, 160 Kvp). After each fraction, cells were grown to ~90% confluence and passaged before receiving subsequent RT doses. The Daoy-WT (DAOY-WT) cell line were subjected to same culture and passaging but were not irradiated.

In-vitro characterization of DAOY-RR

Clonogenic survival: Radioresistance was verified by clonogenic assay. Briefly, the cells were trypsinized into a single-cell suspension and seeded in triplicate at variable densities depending on the radiation dose (higher seeding
density for higher doses) into 6-wells plates. Cells were irradiated 24 hours after plating with a single dose of 0, 2, 4 and 6 Gy (RadSource 2000, 25mA, 160 Kvp). Colonies were fixed and stained 10 days after irradiation with 0.5% crystal violet in ethanol and colonies containing >50 cells were counted using ImageJ (National Institutes of Health, Bethesda, MD). The plating efficiency (PE) and survival fraction (SF) was calculated using the following equations:

\[
(2) \quad PE = \frac{\text{No of colonies formed}}{\text{No of cells seeded}} \times 100\%
\]

\[
(3) \quad SF = \frac{\text{No of colonies formed after treatment}}{\text{No of cells seeded} \times PE}
\]

Survival cures were fitted to a linear-quadratic model (4) by least squares minimization using a \(\frac{1}{\sigma^2}\) weighting using GraphPad Prism 7 to find \(\alpha\beta\) values (GraphPad Software, Inc., La Jolla, CA). The survival cure was used to calculate \(D_{10}\) (dose required to reduce survival to 10%)

\[
(4) \quad SF = e^{-\alpha \cdot \text{dose} - \beta \cdot \text{dose}^2}
\]

**Cell cycle:** Prior to start of experiment culture media was changed to one that is FBS-free for 48 hours to allow cell population to revert to G1/G0-phase. The cells were then switched back to normal culture media for 24 hours prior to start of RT. Cells were harvested, washed with FBS and fixed using cold 70% ethanol for 30 mins. Fixed samples were washed twice in PBS. Cells were then treated with ribonuclease (50 µl of a 100 µg/ml stock of RNase Sigma-Aldrich, St. Louis, MO), and 200 µl propidium iodide (from 50 µg/ml stock solution) for 1 hour. Cell cycle analysis was run with BD LSRII Analysis was performed using the FlowJo software.
Cell cycle was analyzed pre-RT and 24 hours post-RT.

**Tumor implantation**

All animal studies were performed in accordance with guidelines established by Baylor College of Medicine Institutional Animal Care and Use Committees (IACUC). Seven-week-old female severe combined immunodeficient mice (Jackson Lab, Bar Harbor, ME) were implanted in the cerebellum (1 mm right of midline, 1 mm back from lambda and 3 mm deep) with $1 \times 10^6$ Daoy cells (DAOY-WT and DAOY-RR) suspended in 5 µl of matrigel (Sigma-Aldrich, St. Louis, MO) to establish the medulloblastoma tumor model. Animals were monitored for general health and euthanized according to established protocol ($n=48$).

The tumor growth was monitored with bioluminescence imaging by injecting the animals with luciferin (10 µL/g of body weight, Sigma-Aldrich, St. Louis, MO) and imaging them using an IVIS Spectrum *in vivo* imaging system (PerkinElmer, Akron, OH) with automatic exposure settings. At approximately 28 days post implant, animals were treated with radiation, consisting of a fractionated regimen of 12 Gy ($6 \text{Gy} \times 2$ fractions with one day of rest between fractions) using RadSource 200 (25mA, 160 Kvp). Normal tissue was shielded using a 1 inch thick lead piece.

**MR imaging**

**Tumor volume:** Tumor volume was assessed with a $T_1$-weighted post-contrast scan on a Bruker Biospec 9.4 T preclinical MRI scanner (Bruker BioSpin Corp., Billerica, MA). Animals were injected with gadolinium ((Magnevist®, 0.1 µL/g diluted in sterile saline) 10 minutes prior to imaging. During imaging, the animals were kept under inhalation anesthesia with 2% isoflurane-oxygen with respiratory monitoring. $T_1$-weighted images were acquired with a RARE (Rapid Acquisition with Relaxation Enhancement) acquisition which had the following parameters:
TR=1500 ms, TE=8.5 ms, 26 slices, resolution=0.117x0.117x0.5 mm, flip angle=180°, field of view (FOV)=3x3 mm and average=4.

**Relaxometry and Perfusion:**

All MR imaging was performed using a 9.4T MR system (Bruke BioSpin, Rheinstetten, Germany). A RAREVTR_T1 scan was acquired with the following parameters: FOV=3 cm, voxel size = 0.0234 x 0.0234 x 1cm, TE=10.43 ms, TR=100, 138.119, 183.115, 238.026, 308.496, 406.986, 571.837, and 1250 msec, Rare Factor=1, NEX=1. T\(_1\) relaxation time was calculated with built in Image Sequence Analysis and t1 sat function. T\(_2\) relaxation time was calculated with built in Image Sequence Analysis and t2 vtr function. A RAREVTR_T2 scan was acquired with the following parameters: FOV=3 cm, voxel size = 0.0234 x 0.0234 x 1cm, TR=1445.584 msec, TE=10.43, 20.87, 31.30, 41.73, 52.16, 62.60, 73.03, and 83.46 msec, Rare Factor =1, NEX=1.

To determine perfusion rate, arterial spin labeling (ASL), a FAIR_RARE scan was acquired with the following parameters: FOV=3 cm, voxel size = 0.0234 x 0.0234 x 1cm, TE=42.2 msec, TR=18000 msec, Rare Factor=72. Perfusion was calculated with built in Bruker ASL_Perfusion_Processing macro and default settings.

**PET imaging**

Animals were imaged with flurodeoxyglucose (FDG) a glucose analogue pre-RT and 48 hours post-RT. All CT and PET images were acquired with an Inveon scanner (Siemens AG, Knoxville, TN). Prior to imaging, animals were fasted for approximately 12 hours. One hour prior to scanning, each animal received approximately 11.1 MBq (300 μCi) of \(^{18}\)F-FDG (Cyclotope, Houston, TX) through intraperitoneal injection (IP). A respiratory pad was placed under the abdomen of the animal to monitor respiration (Biovet, Newark, NJ) during imaging. Animals
were anesthetized with isoflurane gas (1–3%) mixed with oxygen at a flow rate of 0.5–1 L/minute, and adjusted accordingly to maintain normal breathing rates. A CT scan was acquired with the following specifications: 220 acquired projections covering 220°, a source to detector distance of 312.91 mm and a source to center rotation distance of 183.92. Each projection was 650 ms with x-ray tube voltage and current set at 80 kVp and 500 μA, respectively. A 30 minute PET scan was immediately acquired following CT acquisition. The PET scans were reconstructed using OSEM3D reconstruction method and registered to the CT scan for attenuation correction. PET/CT images were manually registered to the T1-weighted tumor volume scan to segment tumor region on PET/CT scan. Standardized uptake value (SUV) is normalized to the animal’s body weight.

Clinical data

Imaging data was acquired from Texas Children’s Hospital. For each patient conventional scans were used for tumor segmentation which included: T1-weighted, T2-weighted, T1-weighted post contrast and fluid-attenuated inversion recovery (FLAIR) scan. Patients were treated according to standard treatment protocol, SJMB96 and SJMB03, with 5-year survival as the end point for treatment outcome.

Image feature extraction

Image features were extracted using a custom program developed in 2016 MATLAB (The MathWorks Inc., Natick, MA). Using a graphical user interface (GUI), a region of interest can be segmented manually and image features are automatically calculated from the selected region. All the images are normalized and resampled using 3D Slicer (open source software, www.slicer.org) as needed. In this work, the entire tumor region was manually segmented for image feature extraction. Tumor segmentation was manually performed by selecting the tumor border as delineated by the enhancement from the contrast agent. For the GLCM representing
the entire tumor region the texture features were calculated by averaging the GLCM from each image slice.

Once the ROI or tumor region was segmented 33 different image features were extracted automatically by the software. These image features include both first and second order features and were derived from the gray level co-occurrence matrix (GLCM). First order features were derived from the intensity distribution of the pixels from the ROI. Second order image features were derived from the GLCM constructed with 7 different gray levels (8, 16, 32, 64, 128, 256 and 512) which included features from Haralick et al, Soh et al and Clausi et al. The final texture features were extracted from the normalized GLCM from four different offsets ($\alpha = 0^\circ, 45^\circ, 90^\circ$ and $135^\circ$ with symmetry and $\delta = 1$). The entire list of image features can be found in Appendix 1.

**Prediction model**

A classification model was constructed in MATLAB with two classes: WT and RR where the image features were used as inputs. In this study three different classification models were investigated: decision tree (fitctree), random forest (TreeBagger) and support vector machine (fitcecoc with Gaussian kernel). To evaluate the algorithm’s performance and prevent overfitting during the training phase a 10-fold cross validation was performed. All hyperparameters were optimized using Bayesian optimization. Feature importance for the random forest was determined by the summation of changes in risk due to splits for every feature and normalizing by the summation of branching point.

**Statistical analysis**

Statistical analyses were conducted using Prism 7 (GraphPad Software, Inc., La Jolla, CA). Statistical significance was established using a two-tail Student t-test with significance established with $p \leq 0.05$ and correct for multiple comparisons.
using the Turkey method when appropriate. Data contained in this work is reported as mean ± standard deviation.

3.3. Results

Radioresistance phenotype established with repeated irradiations

Radioresistance was verified in the DAOY-RR cell line with clonogenic assay. There was no significant difference in survival fraction at lower doses however survival fraction diverged at higher doses with DAOY-RR cells showing higher cell survival. Experimental data was fitted to a linear-quadratic model. The fitted linear-quadratic curve was used to calculate $D_{10}$ for each cell line. $D_{10}$ was found to be significantly higher in the DAOY-RR cells than DAOY-WT indicating a higher dose is required to achieve 10% survival (Table 2, $p<0.0001$). The survival curve for DAOY-RR cell line had a more linear slope compared to DAOY-WT which seems to indicate more sensitivity to single hit damage (Figure 5).

<table>
<thead>
<tr>
<th></th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$D_{10}$ (Gy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAOY-WT</td>
<td>0.074±0.015</td>
<td>0.070±0.003</td>
<td>4.99±0.03</td>
</tr>
<tr>
<td>DAOY-RR</td>
<td>0.187±0.013</td>
<td>0.019±0.002</td>
<td>7.10±0.07****</td>
</tr>
</tbody>
</table>
Cell cycle analysis show that DAOY-RR cells have significantly more cells in G1 (WT=34.36±0.15%, RR=36.0±0.6%) and G2 phase (WT=16.9±0.9%, RR=18.7±1.8%) compared to DAOY-WT which seems to indicate more cells are in the processing of dividing but not necessarily more active DNA replication. At 24 hours post-RT there was a significant increase in G1 but decrease in G2 and S phase compared to pre-RT for both cells lines. However relative to DAOY-WT population a larger population of the DAOY-RR is still undergoing cell division as indicated by higher percent cell population in the S and G2 phase after RT (Figure 5B).

Figure 5. In-vitro characterization of radioresistant DAOY (DAOY-RR) and wild-type DAOY (DAOY-WT) cell lines. A. Clonogenic survival of DAOY-RR/WT cells exposed to 160 kVP x-ray as a function of dose. Clonogenic assay demonstrate higher survival rates in DAOY-RR cells, solid line is the fitted linear quadratic model. B. Cell cycle distribution of each cell line before and 24 hours post-RT as measured by PI staining. All groups were significantly different. Data is shown as mean±standard deviation, ***p≤0.001.
Radioresistant phenotype is preserved in-vivo

The tumor growth was monitored with bioluminescence imaging where tumor volume is proportional to the maximum photon flux. At approximately 28 days post implant animals were treated with RT. After RT, the animals continued to be monitored with bioluminescence imaging until no signal was detected or until the animal required euthanasia due to tumor burden. Without RT, tumor growth between DAOY-WT and DAOY-RR diverged at around 42 days post-implant where DAOY-RR showed slower tumor growth. Furthermore these tumors did not reach the same tumor volume as DAOY-WT at the terminal time points. This may be due to higher incidences of metastatic disease. With bioluminescence imaging, and later verified with MRI, DAOY-RR tumors were found to be twice as likely to exhibit metastatic disease as indicated by bioluminescence signal in the spinal cord and enhancement in the spinal from MR images. With the administration of RT, DAOY-WT showed complete local tumor control while DAOY-RR tumors continued to progress (Figure 6).
Figure 6. Tumor growth curve. Tumor growth was monitored with bioluminescence imaging where photon flux is proportion to tumor volume. For DAOY-RR tumors, following RT, growth was delayed but resumed growth after 20 days post RT. DAOY-WT showed complete tumor control with no visible signal. DAOY-RR tumors had higher incidences of metastatic disease as highlighted in the figure insert. WT=DAOY-WT, RR=DAOY-RR, data is shown as mean±standard deviation and scale bar=0.5 cm.

Functional changes detected after RT

\(T_1\) and \(T_2\) relaxation time did not differentiate DAOY-WT and DAOY-RR (\(T_1\): WT= 1170±150 msec, RR=1260±140 msec; \(T_2\): WT= 56±9 msec, RR=58±12 msec). There was also no significant changes at 48 hours post-RT in either DAOY-WT or DAOY-RR (\(T_1\): WT= 1190±120 msec, RR=1400±400 msec; \(T_2\): WT= 61±14 msec,
RR=70±20 msec. Tissue characterization with T<sub>1</sub> and T<sub>2</sub> relaxation time could not differentiate the two tumor types and was not affected by RT (Figure 7A,B).

There was a significant decrease in blood perfusion post-RT in DAOY-WT tumors but not DAOY-RR. No significant difference was found between DAOY-WT and DAOY-RR pre-RT (WT= 88±19, RR=60±30 ml/[100gxmin]) which seems to indicate that intrinsically these tumors do not have different vasculature or at least blood perfusion to the tumor is similar. A significant decrease in perfusion after RT in DAOY-WT indicates that the treatment had significantly altered the blood flow through the tumor but DAOY-RR tumors were not affected and showed no significant change (WT= 80±20, RR=60±30 ml/[100gxmin]) (Figure 7C).

Metabolic activity was not significantly different between the two DAOY tumors as measured by <sup>18</sup>F-FDG PET imaging with either mean SUV or max SUV (Table 3). Furthermore RT did not impact the metabolic activity of these tumors.

**Table 3.** Metabolic activity of tumor region as measured by <sup>18</sup>F-FDG PET imaging.

<table>
<thead>
<tr>
<th></th>
<th>WT-Mean SUV</th>
<th>RR-Mean SUV</th>
<th>WT-Max SUV</th>
<th>RR-Max SUV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-RT</td>
<td>1.01±0.13</td>
<td>1.1±0.3</td>
<td>1.3±0.2</td>
<td>1.4±0.3</td>
</tr>
<tr>
<td>Post-RT</td>
<td>0.96±0.14</td>
<td>1.00±0.17</td>
<td>1.3±0.2</td>
<td>1.4±0.3</td>
</tr>
</tbody>
</table>
Figure 7. Functional imaging of DAOY-WT and DAOY-RR pre and post-RT.

Measured tumor’s A. $T_1$ relaxation time, B. $T_2$ relaxation time preRT and 48 hours post-RT. There were no significant differences between $T_1$ and $T_2$ relaxation time between tumors or treatment effects. C. Tumor blood perfusion as measured by ASL pre and 48 hours post-RT. Perfusion decreased in DAOY-WT 48 hours after RT. Data is shown as mean±standard deviation, **$p\leq0.01$. 
Prediction model performance

Three different classification models were investigated: decision tree, random forest and support vector machine. In addition to the three supervised machine learning techniques, we also examined the effect the number of gray levels has on the models’ performances. The number of discrete gray levels dictates the size of the GLCM and therefore impacts the extracted image features. Our results show that the classification model of choice as well as the number of gray levels can impact model performance (Figure 8).
Figure 8. Validation accuracy of each classification model. Validation accuracy of A. Pre-RT classification model performance, and B. Post-RT classification model performance. Each point on the box-whisker plot represents the number of discrete gray level used to construct the GLCM. The box plot represents 25th and 75th percentiles; the bar = min and max; line=median.

The best performing model to predict radioresistance pre-RT was a support vector machine (SVM) with 16 gray levels, which achieved a validation accuracy of 77% (Figure 9A). Sensitivity was 0.81 and 0.74, specificity was 0.71 and 0.81 and precision was 0.76 and 0.78 for WT and RR respectively. Both decision and random forest had poor performance with a validation accuracy of less than 50%. However when the same model was applied to post-RT image features, the performance greatly decreased to achieve a validation accuracy of only 67%. Even though there was a significant decrease in model performance, the overall performance of all the
other classifiers increased when trained with post-RT image features. The decision tree with 32 gray levels had the highest validation accuracy of 93% (Figure 9B). Sensitivity was 0.97 and 0.86, specificity was 0.86 and 0.97 and precision was 0.89 and 0.96 for WT and RR respectively. Interestingly, the decision tree contained only one decision node, image contrast, who’s threshold was 0.761, even though this feature was not statistically different (WT=0.6±0.2, RR=0.97±0.16, when corrected for multiple comparison). This seems to indicate the tumor’s response shortly after RT is more predictive of the long term outcome.

Figure 9. AUC of classification model. A. Model performance of 16 gray level SVM classifier pre-RT as determined by the area under the curve (AUC). B. Model performance of 32 gray level decision tree post-RT as determined by the area under the curve (AUC).
Clinical data

Due to the differences in the image data structure between the preclinical model and clinical data it was not possible to derive meaningful information using the previously established gray levels from the preclinical settings. This was likely due to differences in the distribution of gray level intensities in the data structure or images. To work around this issue, the GLCM size was increased to 512, 1024 and 2048 and for each GLCM size, features were extracted from clinical images and used in the different classifiers (decision tree, random forest and SVM). Figure 10 highlights the four different scans that were used, T1-weighted, T1-weighted post contrast, T2-weighted and fluid attenuated inversion recovery (FLAIR), for tumor region segmentation and feature extraction. Instead of WT and RR classification each image data set was classified as responding or non-responding based on 5-year survival outcome. Similar to the pre-clinical workflow, each classifier was optimized for each scan protocol and GLCM size. Our results show that overall T1-weighted scans were the least informative with T2-weighted and T1-weighted post contrast having the best performance. The use of multiple scan types with a given classifier did not significantly improve the classifier’s performance (Figure 10).

The best performing models are highlighted in Figure 11. Overall these models have similar area under the curve (AUC). The most informative features used in the random forest model consists of area (0.00767), skewness (0.00895), correlation (0.00339), cluster prominence (0.00441) and cluster shade (0.00517).
where the least important feature (no impact on random forest prediction) has a value of 0 and a value of 1 corresponds to a completely deterministic feature.

Two features were found to be significantly different ($p \leq 0.001$), cluster shade (responder$=0.4 \pm 0.9$, non-responder$=-3.1 \pm 1.2$) and cluster prominence (responder$=17 \pm 30$, non-responder$=48 \pm 8$) with a $T_1$-weighted post contrast scan and 512 gray level.
Figure 10. Example of the images used for feature extraction. Each patient was classified as responder or non-responder and extracted image features were passed to different classifiers. Prediction model was optimized for each scan and the three specified gray levels (512, 1024 and 2048). Validation accuracy was calculated for each scan type, given gray level and classifier. Validation accuracy varied across scan type and gray levels within a given classifier where the white arrow points to the best performing model.
Figure 11. Measures of classifier performance. A. Best performing decision tree which consists of cluster prominence (cprom) and kurtosis. B-D, AUC performance for B. T2-weighted scan with 512 gray level using a decision tree, C. T1-weighted post-contrast with 20148 gray levels using a random forest model and D. T1-weighted post contrast scan with 512 gray level using a support vector machine.
3.4. Discussion

In this study we developed a radioresistant medulloblastoma cell-line (DAOY-RR). These cells were used in-vivo as a radioresistant tumor model to identify image features that are indicative of treatment outcome. Image features derived from standard medical images such as a T1-weighted post contrast images were found to be predictive of treatment outcome with up to 74% accuracy. Features derived from scans acquired 48 hours post-RT had an increased accuracy of 93%. The performance of classifiers were found to be dependent on the number of gray levels or size of the GLCM. The preclinical prediction models however could not be directly used on clinical images due to discrepancy in pixel intensity distribution. In spite of this, we used the preclinical model to develop a reproducible radioresistance medulloblastoma tumor model and established a workflow that demonstrates the utility of standard medical images to predict treatment response prior to starting treatment.

We did not find significant differences between T1 and T2 relaxation time of wild-type and radioresistant tumors. This is consistent with previous clinical studies which found that tissue heterogeneity of intracranial tumors leads to an overlap in relaxation time which makes it difficult to differentiate different tumor types. However, T1 relaxation time has been useful in assessing response to bevacizumab in preclinical ovarian cancer models and detection of cirrhosis in the liver which
seems to suggest that the applicability of using tissue relaxation time may be limited to specific tissue types\textsuperscript{73,74}.

In addition we also did not observe significant metabolic changes between DAOY-WT and DAOY-RR or changes after RT. This may be due to poor PET resolution where the average tumor volume is approximately 2.3 mm\textsuperscript{3} while the voxel size of the scanner is 0.39x0.39x0.80 mm. However, \textsuperscript{18}F-FDG PET imaging has been used in other tumor types to access therapy response in the clinic with the majority of these study reporting a decrease in uptake after RT which is associated with favorable outcome\textsuperscript{75,76}. Therefore larger tumor volumes would be needed if PET imaging is be investigated in preclinical models of radioresistance.

Arterial spin labeling or ASL is an imaging technique used widely in the clinic to assess cerebral blood flow with applications ranging from assessing neurodegenerative disease and neurological disorders and distinguishing between true tumor progression and radiation necrosis\textsuperscript{77}. At 48 hours post-RT we observed a significant decrease in perfusion in DAOY-WT tumors but not DAOY-RR. Hong et al. have shown in a rat glioma model that there is a significant decrease in tumor blood flow post-RT\textsuperscript{78}. However the authors also points out that ASL will produce large standard deviations, which we also observed in our study, and that these measurements are sensitive to bulk water intensity which lowers the signal to noise ratio, requiring further normalization\textsuperscript{79}. Further validation is needed to establish the use of ASL and therefore tumor perfusion as a marker of radioresistance.
Predictive performance in the clinical data was not as robust as that observed in the preclinical model. Several factors can contribute to this including difference in the scan acquisition, reconstruction protocols between sets of imaging data and the tumor volume. In the preclinical datasets, image acquisition was highly controlled to prevent noise from being introduced into the model, whereas the clinical data was analyzed retrospectively, and was not acquired in a controlled manner optimized for our study. For example, in the clinical data, wide ranges of tumor volumes were observed whereas the tumor volumes in our study were held constant. While we did not identify tumor volume as one of the primary features indicative of treatment outcome, other image features are often highly correlated with tumor volume, leading to a misattribution of variations in image features to RT responsiveness when the variations in image features are simply being driven by variations in tumor volume.

Previous work have shown that standard medical images can be used to classify pediatric brain tumors with varying degrees of success depending on the image used. Rodriguez Guiterrez et al used a support vector machine classifier with similar features to one proposed in this study had 71% and 85% accuracy on T1-weighted and T2-weighted scans respectively. Perreult et al have shown that standard images can also be used to predict molecular subgroups. However the study was done with trained radiologist and all images were acquired at a single institute. They found that tumor location was one of the most predictive parameter, which suggests that clinical data may also be informative of treatment outcome.
However, to our knowledge there are currently no published study on the use of texture analysis to predict treatment outcome.

While we have shown promising results with the use of texture features to predict treatment outcome, there are a few limitations with this work. The current prediction models are supervised machine learning algorithms which relies on a predetermined class structure and training data sets. Therefore the predictive model will only provide binary classification: WT or RR and with clinical data, responsive and non-responsive, specifically for a medulloblastoma tumor model. If a new input is given that is not a medulloblastoma tumor the model will most likely fail, limiting its application to this particular tumor type. However, this model can be extended to include other brain tumors by specifying new classes into the model and providing new training data.

Another limitation to our study is that preclinical prediction models are not easily translated into the clinic, particularly since this requires the model to be reconstructed when dealing with clinical data. However from this work we have established a workflow that takes any medical image, performs a segmentation and extracts the relevant image features. By testing different classification models with different GLCM size we have found a model that offers superior performance. This seems to indicate that future work with using image features to construct a prediction model should be aware of the constraints of model selection and work towards standardization of image feature extraction.
3.5. Conclusions

From this study we have demonstrated the utility of texture features to predict treatment outcome. The development of a radioresistant tumor model that is validated in both in-vivo and in-vitro studies can also provide an additional tool for researchers to better understand radioresistance. Furthermore these features are derived from standard imaging protocols which can help our workflow become readily available to clinicians without disturbing the current clinical workflow. Early identification of those who will or will not respond to treatment can help identify those who will need more specialized care and those who may benefit more from forgoing radiation and undergoing surgical resection and chemotherapy alone.
Chapter 4

Robustness of Radioresistant Image Predictors: Glioma model

4.1. Introduction

Brain tumors are commonly occurring cancer in pediatric population (approximately 24,000 new cases each year) that are highly aggressive generally have poor treatment outcomes (only 33.2% 5 year survival)\textsuperscript{81}. Management of these tumors usually includes intracranial surgery, radiation therapy and chemotherapy with treatment plans tailored to the diagnosed tumor type. The diagnosis is made based on biopsy in combination with molecular tests on the resected tumor tissue. However for some patients, biopsy or surgical resection is not possible therefore diagnosis relies solely on medical images\textsuperscript{81}. While biopsies are the gold-standard in diagnosis, they come with limitations such as surgical risks, limited spatial and temporal resolution and subjectivity of immunohistochemical scoring\textsuperscript{82,83}. Current advances in imaging technology and analysis has made it possible to establish tumor
phenotype, offering an additional source of information to complement biopsies and an alternative diagnostic tool when biopsies are not possible.

Brain tumors are the second most common malignancy in the pediatric population. The prognosis for glioma is very poor compared to medulloblastoma with the current standard of care which includes surgical resection, radiation therapy and/or chemotherapy, in which 5 year survival is 15 - 35% compared to 70%, respectively. For tumors such as medulloblastoma that are more responsive to treatment, it may not be necessary for young patients to undergo the inherently risky surgical resection. Furthermore, the ability to predict tumor growth can help clinicians make informed decisions in cases where treatment needs to be delayed. For children under the age of three, radiation therapy is delayed due to the long term side effects on the developing brain, therefore establishing the tumor's growth rate can identify critical treatment time points.

Currently in the clinic, imaging is primarily used for anatomical information such as assessing tumor volume and location, determining feasibility of surgical resection and assessing response to treatment. Additional information can be extracted from these images using radiomics by mining the images for quantitative image features that are not intuitively observable such as variance in neighboring pixel values. Recent studies have found that these features can be informative of the tumor's underlying molecular processes and provide valuable diagnostic, prognostic and predictive information. In this work, we focus on texture...
features derived from gray level co-occurrence matrix (GLCM) which are based on the statistical measure of the frequency of occurrence of pairs of pixels with specific intensity and orientation\textsuperscript{67}. These features, however, can be sensitive to image processing which include acquisition, reconstruction protocols and inter-scanner variability\textsuperscript{90–94}. Independent of these systemic variations the values of these features can also be affected by the GLCM size or the number of gray levels which is determined a priori to feature extraction. There have been few studies on the impact of these features from MR images which highlights the need for such preclinical studies where imaging parameters can be controlled and confounding factors can be eliminated\textsuperscript{95}.

The aim of this work is to classify brain tumor type and predict tumor growth rate using texture features from T1-weighted post contrast MR scans in a preclinical model. Tumor regions were segmented using in-house software with GLCMs constructed for a single tumor slice and entire tumor volume. We investigated the sensitivity of texture features values to different GLCM sizes and how this affects the performance of different classifiers. We hypothesized that due to the heterogeneity of tumors, 3D volumetric data would have the highest accuracy and that a larger GLCM size will be more informative of the tumor region. To further investigate the use of these texture features to predict growth rate, a shallow neural network was constructed to predict the $\alpha\beta$ value of an exponential function using texture features derived from diagnostic images. Using machine learning we will determine whether radiomics approaches have the potential to classify tumor type and predict tumor
growth rate noninvasively and therefore allow clinicians to make more informed decisions using standard medical images.

Figure 12. T1-weighted post contrast MR scan of A. GL261, mouse glioma, B. U87, human glioma, C. Daoy, human medulloblastoma D. Schematic of workflow for feature extraction. Images are imported into a custom program for visualization and segmentation of the tumor region. The selected region used to extract 1st and 2nd order image features. These features are then used to construct the three different supervised classifiers: Decision Tree, Random Forest, and Support Vector Machine.
4.2. Methods

All animal studies were performed in accordance with guidelines established by Baylor College of Medicine Institutional Animal Care and Use Committees (IACUC). Severe combined immunodeficient mice (Jackson Lab, Bar Harbor, ME) were implanted in the cerebellum with $1 \times 10^6$ Daoy cells (ATCC, Manassas, VA) suspended in matrigel (Sigma-Aldrich, St. Louis, MO) to establish the medulloblastoma tumor model. Glioma models were established in the caudate putamen. The human glioma model was established in severe combined immunodeficient mice with $1 \times 10^6$ U87 (ATCC, Manassas, VA) cells suspended in matrigel. The mouse glioma model was established in C57BL/6 Albino with $5 \times 10^5$ cells in matrigel. Animals were monitored for general health and euthanized according the established protocol. Each experimental group included 7-10 animals.

Texture feature extraction

Animals were imaged either weekly or biweekly based on tumor type, with $T_1$-weighted (TR=1500 ms, TE=8.5 ms, matrix size 256 x 256, pixel spacing 0.117 x 0.117 mm, slice thickness 0.5 mm) post contrast scan on a 9.4T magnet (Bruker BioSpin). Gadolinium (Magnevist®, 0.1 µL/g diluted in sterile saline) was administered intravenously 10 minutes prior to start of image acquisition. The acquired images were used for texture feature extraction (n=87).

Image features were extracted using a custom program developed in 2016 MATLAB (The MathWorks Inc., Natick, MA). The program allows the users to import the image files and manually select a region of interest (ROI) using a graphical user interface (GUI). All the images are normalized and resampled using 3D Slicer as needed. Tumor region segmentation was performed on the central, middle, edge and entire tumor region. The central slice was defined as the slice with the largest cross section, the edge of the tumor was designated as the second to last slice the
tumor was visible and the middle tumor region was defined as halfway between the center and tumor edge. For the GLCM representing the entire tumor region the texture features were calculated by averaging the GLCM from each image slice. Tumor segmentation was manually performed by selecting tumor border as delineated by the enhancement from contrast agent.

Once the ROI or tumor region was segmented 33 different image features are automatically extracted. These image features include both first and second order features and were derived from the gray level co-occurrence matrix (GLCM). First order features were derived from the intensity distribution of the pixels from the ROI. Second order image features were derived from the GLCM constructed with 10 different gray levels (8, 16, 24, 32, 48, 64, 98, 256, 512) which included features from Haralick et al, Soh et al and Clausi et al\textsuperscript{67,70,71}. The final texture features were extracted from the normalized GLCM from four different offsets (0\degree, 45\degree, 90\degree and 135\degree with symmetry and closest neighbor). The list of image features can be found in Appendix A.

**Tumor type classification model**

The classifications models were constructed in MATLAB with three classes: GL261, U87 and Daoy and the image features as inputs. In this study three different classification models were investigated: decision tree (fitctree), random forest (TreeBagger) and support vector machine (fitcecoc with Gaussian kernel). To evaluate each algorithm's performance and prevent overfitting during the training phase a 10-fold cross validation was performed. All hyperparameters were optimized using Bayesian optimization. Feature importance for random forest was determined by the summation of changes in error due to node removal and normalizing by the number of branching points.
Tumor growth rate prediction

Tumor growth curves from each tumor type were first fitted to a one-term exponential:

\[ \text{tumor volume} = \alpha e^{\beta \cdot \text{time}} \]  \hspace{1cm} (5)

The \( \alpha \beta \) values were fitted using the Trust-Region algorithm with default setting using MATLAB. The values found from the fit were used as the target value for the neural network.

The two layer neural network used consisting of 33 hidden neurons with a sigmoid transfer function in the hidden layer and linear transfer function in the output. The network was trained with Levenberg-Marquardt backpropagation algorithm where 60% of the data was used for training, 35% used for validation and 5% used for testing. Training and testing samples were divided to have similar distribution and equal representation of all three tumor types.

Statistical analysis and model performance evaluation

Statistical analysis was performed using GraphPad Prism (GraphPad Software, La Jolla, CA). Model performance was assessed with the following metrics: accuracy, specificity, sensitivity and F-score. TP = true positive, TN = true negative, FP = false positive, FN = false negative:

\[ \text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \]  \hspace{1cm} (6)

\[ \text{Specificity} = \frac{TN}{FP + TN} \]  \hspace{1cm} (7)
\begin{align}
(8) \text{Sensitivity} &= \frac{TP}{TP + FN} \\
(9) \text{Fscore} &= \frac{2 \cdot TP}{2 \cdot TP + FP + FN}
\end{align}

The metrics were computed for each individual tumor class from the confusion matrix. Analysis of variance (ANOVA) was used to compare multiple means with statistical significance set at \( p < 0.05 \) for each image feature and correction for multiple comparisons using a Bonferroni correction.

### 4.3. Results

To construct our classification and prediction models, texture features were first extracted from the tumor region using in-house MATLAB software for three different types of tumors: GL261 (mouse glioma), U87 (human glioma) and Daoy (human medulloblastoma). The graphical user interface (GUI) allow users to import in DICOM files either in batch or single image slice and manual segmentation of the region of interest (ROI). Once segmentation is completed the program calculates first and second order image features based on the specified GLCM size which is dictated by the number of specified gray levels. The extracted image features from the T1-weighted post contrast scans for each tumor type are used as inputs to train the classification and prediction models (Figure 12). A separate cohort was used for testing.
Comparison of our radioresistance medulloblastoma (DAOY-RR) with the established glioma model reveal one imaging feature that were specific to treatment response. Cluster prominence was significantly higher (p≤0.05) in U87 and DAOY-RR tumors compared to DAOY-WT (Table 4). We also identified one feature that differentiated (p≥0.05) medulloblastoma from glioma tumors (Table 4). With functional imaging, there were significant functional difference between medulloblastoma and gliomas as demonstrated by T₁ relaxation time (Figure 13). There were however no differences found in T₂ which seems to suggest difference in lipid content but not water composition that differentiate these tumors.

**Table 4.** Features that differentiated tumor types, GLCM with 512 gray levels.

<table>
<thead>
<tr>
<th>Image Feature</th>
<th>DAOY-WT</th>
<th>DAOY-RR</th>
<th>U87</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster Prominence</td>
<td>4.7±3.3x10⁵</td>
<td>5.1±4.3x10⁵</td>
<td>5.6±4.8x10⁵</td>
</tr>
<tr>
<td>Sum of variance</td>
<td>4.4±0.2x10⁵</td>
<td>4.4±0.3x10⁵</td>
<td>5.3±0.5x10⁴</td>
</tr>
</tbody>
</table>
Figure 13. **T₁ and T₂ relaxation time of glioma and medulloblastoma tumors.** For U87 and Daoy tumors **A.** T₁ relaxation time and **B.** T₂ relaxation times was measured at similar tumor volume. T₁ relaxation time is able to differentiate tumor type in which Daoy tumors have a longer relaxation time. Box plot represents 25th and 75th percentiles; bar = min and max values; line=median, ***p≤0.001.

**Tumor Type Classification**

Overall the three different classification models had similar performances. **Figure 14** show the validation accuracy of the different GLCM sizes for each classifier and tumor region. We observed a trend of increasing validation accuracy with the number of gray levels or GLCM size (see Appendix B) however it was not a proportional increase. The selection of the tumor region used for GLCM construction also impacted the classifier’s accuracy. Using the entire tumor volume resulted in higher accuracy compared to a single image slice. The edge of the tumor was the least predictive while the central and middle regions were comparable (see Appendix B).
Figure 14. Validation accuracy of each classification model of given gray level. Model performance as evaluated by validation accuracy for different tumor regions using A. decision tree, B. Random Forest and C. Support Vector Machine. Model performance was consistently poor at the edge of the tumor and high when using the entire tumor volume. D. Receiver operator curve of the best performing classifier for each tumor type, 512 gray levels with random forest. Box plot represents 25th and 75th percentiles; bar = min and max values; each point is a different number of gray levels.
The best performing classification model was using random forest with 512 gray levels which achieved 84% validation accuracy. With the same model using the central, middle and edge tumor regions the validation accuracies were 74%, 74% and 52% respectively. The model resulted in an area under the curve (AUC)=0.92, 0.88 and 0.85 for GL261, U87 and Daoy respectively (Figure 14). While there was high specificity and sensitivity for both GL261 and U87 tumor classifications, specificity was only 0.75 for Daoy tumors (Table 5-7). Even though 256 gray levels GLCM had similar accuracy the classifier performance was not as uniform for all tumor types. Therefore all subsequent analysis were performed on features extracted from the GLCM with 512 gray levels. We've also identified the four texture features dependent on GLCM size: autocorrelation, cluster prominence, sum of square and sum variance (p<0.01-0.0001) (see Appendix B).

Table 5. Classification model performance for GL261 in different tumor regions.

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>Center</th>
<th>Middle</th>
<th>Edge</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC</td>
<td>0.923</td>
<td>0.909</td>
<td>0.960</td>
<td>0.732</td>
</tr>
<tr>
<td>Accuracy</td>
<td>0.90</td>
<td>0.849</td>
<td>0.882</td>
<td>0.725</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.903</td>
<td>0.903</td>
<td>0.770</td>
<td>0.821</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.911</td>
<td>0.810</td>
<td>0.952</td>
<td>0.659</td>
</tr>
<tr>
<td>F-Score</td>
<td>0.875</td>
<td>0.778</td>
<td>0.909</td>
<td>0.622</td>
</tr>
</tbody>
</table>
Table 6. Classification model performance for U87 in different tumor regions.

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>Center</th>
<th>Middle</th>
<th>Edge</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC</td>
<td>0.879</td>
<td>0.842</td>
<td>0.813</td>
<td>0.708</td>
</tr>
<tr>
<td>Accuracy</td>
<td>0.895</td>
<td>0.849</td>
<td>0.765</td>
<td>0.695</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.857</td>
<td>0.65</td>
<td>0.769</td>
<td>0.263</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.909</td>
<td>0.925</td>
<td>0.762</td>
<td>0.86</td>
</tr>
<tr>
<td>F-Score</td>
<td>0.783</td>
<td>0.765</td>
<td>0.667</td>
<td>0.417</td>
</tr>
</tbody>
</table>

Table 7. Classification model performance for Daoy in different tumor regions.

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>Center</th>
<th>Middle</th>
<th>Edge</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC</td>
<td>0.848</td>
<td>0.775</td>
<td>0.865</td>
<td>0.565</td>
</tr>
<tr>
<td>Accuracy</td>
<td>0.882</td>
<td>0.781</td>
<td>0.765</td>
<td>0.623</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.75</td>
<td>0.591</td>
<td>0.5</td>
<td>0.364</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.942</td>
<td>0.863</td>
<td>0.846</td>
<td>0.745</td>
</tr>
<tr>
<td>F-Score</td>
<td>0.857</td>
<td>0.65</td>
<td>0.5</td>
<td>0.4</td>
</tr>
</tbody>
</table>

Random forest is an ensemble learning technique which grows many decision trees where the final class is determined by a majority “vote” from all the decision trees. Analysis of the estimated feature importance reveals skewness (0.91) to be the most important feature followed by cluster shade (0.84), kurtosis (0.55), median intensity (0.40), mean intensity (0.30), sum average (0.30), sum variance (0.289), autocorrelation (0.26), max probability (0.24) and entropy (0.23), where 0 represents the smallest possible importance and higher values have greater importance (Figure 15).
Figure 15. Importance of features in tumor classification. A. Feature importance was estimated for 512 gray level random forest model using the entire tumor volume. Top 10 most important features included both 1st and 2nd order texture features. B. Cluster prominence was calculated for each tumor type. Cluster prominence was significantly difference between medulloblastoma and glioma, both U87 and GL261 tumors. Box plot represents 25th and 75th percentiles; bar = min and max values; line=median, ***p≤0.001, ****p≤0.0001.
Direct comparison of the image features derived from the GLCM with 512 gray levels showed cluster prominence to be the only feature that differentiates glioma tumors from medulloblastoma. Cluster prominence which measures the asymmetry of the GLCM was significantly lower in both glioma models (p<0.001 to GL261, p<0.0001 to U87) compared to medulloblastoma (GL261=(1.3 ± 1.7) x 10^6, U87=(1.2 ± 1.6) x 10^6, Daoy=(1.6 ± 1.1) x 10^6) (Figure 15). This indicates that gliomas demonstrated a lower variance in gray levels.

**Growth Curve Prediction**

Using a shallow two layer neural network, the tumor growth curve could be established with mean error squared of 16% (Figure 16A). The neural network was trained to predict the αβ value for an exponential growth model. The inputs used to train the network were image features from similar tumor volumes and αβ value from the fitted exponential of the experimental data (Figure 16B). The experimental data showed U87 to be highly aggressive with a rapid growth rate (α=1.2±0.8, β=0.21±0.03) and Daoy having a slower growth rate (α=0.9±0.5, β=0.08±0.04) (Figure 16C). GL261 tumors had higher variance in tumor growth compared to other tumor types (α=0.7±0.9, β=0.15±0.04) (Figure 16D). Overall, the neural network was able to predict tumor growth more accurately for U87 and Daoy tumors than GL261.
Figure 16. Tumor growth prediction model. A. A two-layer feedforward network consisting of 33 hidden layers with a sigmoid transfer function in the hidden layer and linear transfer function in the output layer was used to predict tumor growth. B. Growth curve of the three different tumor types with the experimental data (in black circles), the fitted exponential growth model (blue dash line) and the predicted growth curve from the neural network (red line). Prediction was more accurate for U87 and Daoy tumors. C.-D. Growth curve of each tumor type as measured by T1-weighted post contrast images. Growth was more heterogeneous in GL261 tumors than U87 and Daoy. Data is shown as mean±standard deviation.
4.4. Discussion

In this study we have demonstrated that by using texture analysis, standard medical images can be used to classify tumors and predict tumor growth rate. Furthermore we have shown that the size of the GLCM or number of gray levels impacts the performance of the classification models and identified a classifier that significantly outperformed all others. The best performing classifier was a random forest model with texture features derived from the 512 gray level GLCM that used the entire tumor volume. These texture features can be used to predict tumor progression using a shallow 2 layer neural network. Assuming an exponential growth curve, the neural network is able to predict the αβ value using only the diagnostic scan. These classification and prediction models can provide an additional tool for tumor diagnosis and more personalized treatment planning without additional burden to the patient.

Previous works have shown that image features can be used for brain tumor classification and grading96-99. However, these studies used different classification models and parameters and different scan sequences with varying degrees of success100,101. Taking a more reductionist approach, our study is focused on the robustness of texture features for different GLCM sizes and how this in turn affects different classifiers’ performances. A previous study has shown that extending the distance of neighbors does not significantly affect the GLCM102 therefore this parameter was not included in this study. We have identified four features which
included autocorrelation, cluster prominence, sum of squares and sum variance that were dependent on the number of gray levels while the other texture features were more robust. Cluster prominence, which is a measure of the GLCM asymmetry, was the only feature that was significantly lower in the glioma (GL261 and U87) tumors than medulloblastoma. This finding agrees partially with Brynolfsson’s et al work using ADC (apparent diffusion coefficient) maps of glioma and prostate data\textsuperscript{103}. However, this study identified more features being influenced by GLCM size and noted a greater effect of GLCM size on the texture feature value. This seems to indicate that the influence of the GLCM size on texture feature value may vary between image acquisition protocols. Therefore to have meaningful comparison across acquisition protocols the GLCM size should be standardized and not part of the optimization process.

Classifier performance was impacted by the GLCM size but the selection of the ROI for the derivation of the GLCM had the greatest impact on performance. Using the entire tumor volume resulted in higher accuracy than using a single image slice alone. This may be due to an increase in counting statistics and less sparse GLCMs which allows for extraction of meaningful features. This was the case at the edge of the tumor where the ROI was small which resulted in poor performance for all three classifiers. Furthermore using the entire tumor region would better take into account the tumor's heterogeneity.
To our knowledge this is the first demonstration of the use of texture features to predict tumor growth using a neural network. While there are many classical models of tumor growth we chose to fit the experimental data to an exponential growth curve since this is a simple model with only two parameters and provided a good fit\textsuperscript{104}. Overall the neural network is able to predict the $\alpha\beta$ with better performance for U87 and Daoy tumors. The prediction model was best for tumors that had tighter growth curve which helps reinforce the network. The high variance in the GL261 tumor growth curve may be due to immunological factors in the animal model. Since GL261 is a mouse glioma, model, cells are implanted into an immunocompetent animal, potentially leading to an immune response that could delay tumor growth to varying degrees.

A limitation of our work is the use of preclinical brain tumor models. Since we are using supervised machine learning algorithms, the classifier relies heavily on the training dataset. Preclinical models can generate these datasets in a timely manner while controlling for confounding factors such as acquisition and reconstruction protocol which is not always readily available with clinical data sets. Furthermore, with a preclinical model, we can track tumor progression and establish the growth curve which is not always possible with patient data. This work is a proof of concept that demonstrates that texture features can not only classify tumor type with high accuracy but also, using neural networks, map out tumor progression.
These findings demonstrate that the image features extracted from standard medical images have the ability to make diagnoses and even predict tumor growth rate. For patients who are not eligible for a biopsy or tumor resection this is an alternative source of information that can help clinicians make informed treatment plans. Furthermore for patients who may not be immediately eligible for other treatments such as radiation therapy, mapping out the growth curve of these tumors can help clinicians identify critical time points when planning the course of treatment. Additionally, understanding the growth of these tumors can help clinicians differentiate between true tumor progression versus metastatic or pseudo-progression after radiation therapy. More importantly this study adds to the previous body of work on the impact of GLCM size on texture feature value.
4.5. Conclusions

Features derived from standard-of-care images can be used to classify tumor type and map tumor growth using machine learning algorithms. The number of gray levels in the construction of the GLCM influenced the performance of the predictive models. A 512 gray level GLCM with random forest classification model yielded an accuracy of 84% based on tenfold cross-validation. These results are promising in achieving a noninvasive marker for tumor type classification. These texture features are also found to be informative of tumor growth. Using a two layer neural network the αβ values were predicted from the image features. The neural network had mean squared error of 16.02% with better performance for U87 and Daoy tumors. The performance of these models can be greatly improved with the addition of new data sets since both the random forest and neural network relies heavily on training data sets. Finally, standardization of feature extraction and exploration of deep learning techniques can contribute to a more accurate prediction of tumor type and progression with standardized GLCM that can allow for meaningful comparison of texture features.
Chapter 5

Imaging Markers of Radiation-Induced Brain Injury

5.1. Introduction

Brain tumors are the second most common childhood cancer, and are often treated with radiation therapy (RT) in combination with surgical resection and adjuvant chemotherapy, achieving a 5-year survival of 70%\(^2,14\). While RT is highly effective in providing local tumor control, most survivors will experience acute and long term side effects due to radiation-induced brain injury since the field of irradiation also includes healthy normal brain tissue. Radiation-induced brain injury can lead to damage to the microvascular network, demyelination and inflammatory response with activation of astrocytes and microglia\(^{50,105,106}\). Acute molecular and cellular changes such as neuroinflammation can lead to long term side effects which may not fully manifest for months or years after treatment\(^{107-109}\). Survivors of childhood cancer experience reduced cognitive processing speed, deficiency in working memory and attention\(^{107,110}\). Currently, there are no established
biomarkers to determine who may be susceptible to radiation–induced brain injury or how these injuries will progress to affect cognitive performance.

Medical imaging techniques such as magnetic resonance (MR) imaging and positron emission tomography (PET) are viable techniques that can be used to characterize the metabolic, anatomical and functional changes in the brain in response to radiation-induced injuries. Furthermore these techniques can be used to monitor the progression thus establish links between changes in cognitive performances to functional changes in the brain. Previous pilot studies in a mouse model have shown that functional imaging can be used to detect acute changes, 48 hours post-RT, in normal tissue after (Figure 17). Magnetic resonance imaging (MRI) techniques, such as diffusion tensor imaging (DTI) have been successfully used in the clinic to detect imaging changes related to cognitive deficits\textsuperscript{111,112}. Fractional anisotropy (FA) values, reflecting isotropic diffusion, have been used as markers of inflammation, white matter volume and to determine the integrity of neuronal myelination which can be correlative of cognitive performance\textsuperscript{113,114}. DTI data can also be informative of the trajectory of white matter fibers in three-dimensional space using tractography. Using this technique, the brain network or connectome can be modeled to provide insight into the brain’s organization\textsuperscript{115}.

PET imaging with \textsuperscript{18}F-fluorodeoxy-glucose (\textsuperscript{18}F-FDG) is commonly used in the clinic for the detection and staging of cancer\textsuperscript{116,117}, as the \textsuperscript{18}F-FDG, a glucose analogue, accumulates in tissues with high metabolic activity such as solid tumors and sites
with inflammation\textsuperscript{118-121}. Thus, this imaging technique can be used to localize and quantify any neuroinflammation response at these late time points. The neuroinflammation response is thought to lead to long term cognitive changes.

\textbf{Figure 17. Functional imaging changes at 48 hours post-RT.} \textbf{A.} Normal tissue blood perfusion was measured at 48 hours post-RT. Treatment lead to a significant reduction in blood perfusion. \textbf{B.} Measurement of glucose uptake in normal tissue region 48 hours post-RT using \textsuperscript{18}F-FDG PET imaging. RT did not alter normal tissue glucose metabolism as detected by this imaging technique. Data is shown as mean±standard deviation, **p≤0.01.
The aim of this study is to identify imaging markers that will provide further insight into late-term radiation-induced brain injury and identify imaging surrogates for cognitive outcome using MR and PET imaging techniques. Young adolescent rats were irradiated and cognitive performance was evaluated using the 5-Choice Serial Reaction Time Task (5-CSRTT) 12 months post-RT. Imaging data were acquired to investigate structural, connective network and inflammatory changes in the brain due to RT and how these changes may drive cognitive outcome.

5.1. Methods

Establishing radiation-induced brain injury model: late RT effects

All animal experiments were approved by Baylor College of Medicine Institutional Animal Care and Use Committee (IACUC). 20 Male Sprague Dawley rats (Harlan Laboratories Inc, Indianapolis, IN) were singly housed in a reversed 12-hour light/dark cycle (lights on from 8:00 PM-8:00AM). At 28 days of age animals were divided into two groups: Sham and RT. Sham animals were anesthetized for the same amount of time but received no RT. RT animals received fractionated whole brain irradiation of 30 Gy (6Gy x 5days) from behind the eyes and to the back of the ears using a RadSource 2000 X-ray irradiator (150kVp, 25mA, Rad Source Technologies, Inc., Suwanee, GA). The rest of the body was shielded using a 1 inch lead piece.
Figure 18. Experimental time line for radiation-induced brain injury model. Animals receive RT at 28 days of age. At 3 months post-RT 5-CSRTT training starts (shaded box) and lasts for 7 weeks (5 days of training per week). Animals are tested at 12 months post-RT at final stage reached after training and Stage 7. All imaging data was acquired after completion of behavioral testing.

Prior to start of cognitive training/testing animals were food restricted to maintain 85-90% of their free-feeding weight to provide incentive for participation. One week prior to start of cognitive training animals were handled for approximately 5 minutes each for 3 days throughout the week to acclimate to handling and given chocolate pellets to acclimate to the reward offered during cognitive training/testing.

**Cognitive Testing**

Cognition was measured at 12 months post RT using 5-choice serial reaction time task (5-CSRTT). Training was completed with software and equipment from Noldus Information Technology (Lessburg VA) and Med Associates (St. Albans, VT). At 3 months post RT, animals were trained to determine their cognitive baseline. During training animals were first habituated to 5 CSRTT testing equipment with chocolate pellets (form of reward). The testing equipment consists of 5 chamber aperture which is illuminated individually during testing and food magazine where
chocolate pellets are dispensed from when the animal correctly pokes the illuminated aperture. Animals started training when greater than 90% of the pellets were consumed during habituation. Training consists of 35 daily sessions following the criteria established by Bari et al\textsuperscript{122}. The objective of the task is for the subjects to nose-poke one of five apertures in which a stimulus light is shown. During the training phase, once a specified set of criteria is met in a particular stage (accuracy, number of correct trials and omissions), the animal advances to the next stage. Testing stages increases in complexity, by adjustment of the stimulus duration, inter-trial interval, and limited hold time. The following metrics were measured to determine animal’s cognitive performance: percent accuracy, percent correct, premature, and omission\textsuperscript{123}.

Upon completion of the training period the final stage and average stage reached for each group (Sham and RT) was recorded. At approximately 12 months post-RT animals were tested at the final stage reached at the end of the training period as well the average stage reached by the Sham cohort (Stage 7). Testing was 5 days for the final stage reached followed by 5 days of testing at Stage 7.

**Imaging – Diffusion tensor imaging**

After completion of cognitive testing and in-vivo imaging animals were euthanized for *ex-vivo* DTI MRI scans. Samples were prepared according to the method described by Tyzska et al\textsuperscript{124}. Briefly, animals were perfused with heparin-saline followed by 10% buffered formalin. Excess tissue was removed from the skull and the sample was fixed overnight in 10% buffered formalin and transferred to 0.01% sodium azide in saline, and rocked at 4°C for 7 days. Tissue was then transferred to a 5mM Magnevist\textsuperscript{®} (Bayer Healthcare Pharmaceuticals, Berlin, Germany) and 0.01% sodium azide solution and rocked for another 21 days at 4°C. Prior to imaging, the tissue were removed from 4°C and allowed to equilibrate to room temperature.
Brains were imaged using 20 distinct gradient directions on a 9.4T Bruker Biospec MRI scanner (Bruke BioSpin, Rheinstetten, Germany). All DTI processing and connectome analysis was performed as described by Sahnoune et al. Briefly, ROIEditor, DTI studio, template maps from Duke Center for In vivo Microscopy, DiffeoMap, AIR algorithm and LDDMM were used for alignment, segmentation and calculation. DTI studio (www.mristudio.org) was used to generate various volume maps and calculate average FA value. Masked tensors were used for fiber tracking where cutoffs for fiber initiation and fiber continuation were set to FA values of 0.3/0.3 with a 70° angle cutoff. Total fiber number, average fiber length and maximal fiber length calculated. Alignment of the brains was done to template maps obtained from the Duke Center for In Vivo Microscopy (http://www.civm.duhs.duke.edu/SharedData/DataSupplements.htm).

Connectomes were constructed for each animal with N = 132 nodes, network degree of $E = \text{number of edges}$ and a network density of $D = E/[(N \times (N-1))/2]$ representing the fraction of present connections to all possible connections. The following connectome properties were calculated using graph theoretical analysis as implemented in Brain Connectivity Toolbox: normalized clustering coefficient (clustering), characteristic path length (path length), small-worldness (SW) index, global efficiency, mean local efficiency, modularity and transitivity. Path length and clustering coefficient were normalized using 100 benchmark random networks.

**Imaging – T1-weighted/T2-weighted/Proton Density scans**

Animals were imaged on a 9.4T Bruker Biospec MRI scanner (Bruke BioSpin, Rheinstetten, Germany) at 12 months post-RT. Animals were anesthetized with isoflurane gas (1–3%) mixed with oxygen at a flow rate of 0.5–1 L/minute, and adjusted accordingly during imaging to maintain normal breathing rates. A T1-weighted, T2-weighted and proton density scan was acquired for each animal with axial slice orientation and 256 matrix size. A T1-weighted scan was acquired with
the following parameters: slice thickness/interslice distance=1 mm, TR=1500 ms, TE=8.5 ms, NEX=1, Rare Factor=4, Resolution=0.0156 cm/pixel. A T2- weighted scan was acquired with the following parameters: slice thickness/interslice=1 mm, TR=3271.7 ms, TE=22 ms, NEX=4, Rare Factor=4, Resolution=0.0156 cm/pixel. A proton density scan was acquired with the following parameters: slice thickness/interslice=1 mm, TR=3682 ms, TE=13 ms, NEX=2, Resolution=0.0156 cm/pixel. White matter which was calculated as percent of the intracranial volume (ICV), gray matter and cerebral spinal fluid was analyzed using an automated hybrid neural network segmentation and classification as described by Reddick et al\textsuperscript{136}.

**Imaging – Positron emission tomography**

At 12 months post RT animals were imaged with \textsuperscript{18}F-Flurodeoxy-glucose (\textsuperscript{18}F-FDG). All CT and PET images were acquired on an Inveon scanner (Siemens AG, Knoxville, TN). Prior to imaging, animals were fasted approximately 12 hours. One hour prior to scanning, each animal received 12.58 MBq (340 μCi) of \textsuperscript{18}F-FDG (Cyclotope, Houston, TX) through intravascular (IV) injection. A respiratory pad was placed under the abdomen of the animal to monitor respiration (Biovet, Newark, NJ). Animals were anesthetized with isoflurane gas (1–3%) mixed with oxygen at a flow rate of 0.5–1 L/minute, and adjusted accordingly during imaging to maintain normal breathing rates. A CT scan was acquired with the following parameters: 220 acquired projections covering 220°, a source to detector distance of 312.91 mm and a source to center rotation distance of 183.92. Each projection was 650 ms with x-ray tube voltage and current set at 80 kVp and 500 μA, respectively. A 30 minute PET scan was immediately acquired afterward. The PET scans were reconstructed using OSEM3D reconstruction method and registered to the CT scan for attenuation correction. Using the acquired CT scan the entire brain region is segmented using Inveon Research Workspace (Siemens AG, Knoxville, TN). Uptake of radioisotope was measured using the standardize uptake value (SUV) and max SUV normalized to animal's body weight.
Immunohistochemistry

Whole brains were excised, serially sliced into 2 mm coronal sections, and embedded in paraffin. Slides were deparaffinized and rehydrated in a standard series of xylene and alcohol. Heat induced antigen retrieval was performed for 30 min at 95°C in sodium citrate buffer pH 6.0 in a microwave. Indirect fluorescent immunohistochemical staining was performed on 5 µm sections using antibodies for rabbit anti-TNFα (1:200, Abcam, ab9635), mouse anti-MBP (1/500, Abcam, ab62631), goat anti-GFAP (1/500, Abcam, ab53554), and goat anti-IBA-1 (1/500, Abcam, ab5076). Fluorescent secondary antibodies conjugated to 488, 555, and 647 nm fluorophores (Alexafluor, Invitrogen) were appropriately matched to the primary antibodies, diluted 1/500, and incubated in the dark for 2 h. Cell nuclei were visualized by DAPI (1/500). Slides were imaged using a Nikon A1 confocal microscope (Tokyo, Japan) at 20X and analyzed using ImageJ software (v.1.50i, NIH, Bethesda, MD). Every TNFα-, GFAP- and IBA-1-expressing cell was imaged within three representative images within the peri-corpus callosum by myelin basic protein (MBP) immunostaining and normalized to total cells by DAPI presence. All corpus callosum regions were optimally imaged at the body, excluding the cingulum bundle. Myelin thickness was measured at the corpus callosum plus the alveolus by Image J, with distances normalized to pixel size. All cells were counted manually per image per region.

Statistical analysis

Statistical analyses were conducted using Prism 7 (GraphPad Software, Inc., La Jolla, CA). Statistical significance was established using a two-tail Student t-test with significance established with p≤0.05 and correct for multiple comparisons using the Holm-Sidak method when appropriate. 5-CSRTT testing at 12 months post-RT cognitive performance was averaged across 5 testing days (at final stage reached and Stage 7). Data is reported as mean±standard deviation. Correlation analysis of image feature and cognitive performance was performed using Pearson’s
correlation with significance established at \( p \leq 0.05 \). P-value for Pearson’s correlation coefficient uses the t-distribution. Connectome statistics were performed according to network based analysis\(^{137}\).

**5.2. Results**

**Brain growth**

Brain growth was stunted early on after RT, persisting up to 12 months post-RT. At 3 months post-RT, Sham brains were significantly larger than RT brains (Sham=1.90±0.05g, RT=1.68±0.09g, \( p=0.004 \)). While Sham brains showed growth from 3 to 12 months (\( p=0.016 \)) RT brains remained significantly smaller (Sham=2.03±0.06g, RT=1.740±0.104g, \( p=0.001 \)) (Figure 19). At 3 months post-RT a ratio of brain’s length and width was determined. Length was measured by the distance from tip of the olfactory bulb to the back of the cerebellum and width was measured at the widest part of the brain. Brain growth was found to be stunted in only one direction, length-wise. At 3 months (Sham=0.754±0.012, RT=0.815±0.010) and 12 months (Sham=0.75±0.02, RT=0.8±0.0.3) post-RT, Sham brains were significantly longer but not wider than RT brains (\( p=0.00024 \) at 3 months and \( p=0.0087 \) at 12 months) (Figure 19).
Figure 19. RT stunts brain growth. A. Image of Sham and RT brains at 12 months post-RT. Brain of Sham animals are visibly larger compared to RT animals, scale bar=1 cm. B. Brain weight was measured at 3 month and 12 months post-RT. Differences in weight is observed early on and persists up to 12 months post-RT with no growth in RT brains. C. Brain dimensions were measured at 3 months and 12 months post-RT, width is measured at the widest section of the brain and height is measured rostral to caudal from the olfactory bulb to the end of the cerebellum. RT brains were significantly shorter while the width of the brain is preserved with no significant changes over time. Box-and-whisker plot: box extends from the 25th to 75th percentiles, median value marked by line and whiskers extend from min to max. *p≤0.05, **p≤0.01, ***p≤0.001.
Unidirectional stunt in growth was largely due to the heterogeneous response of the different brain regions. Volumetric analysis of specific brain regions using DTI, showed that 9 of the 26 brain regions analyzed were significantly smaller in RT brains than Sham brains (Figure 20, Appendix C). Cerebellum was the only region showing increase after RT (Sham=12.4±0.6%, RT=13.3±0.3%, p=0.013). The affected brain regions are not localized to any specific area of the brain.

**Figure 20.** Volumetric data from DTI scans demonstrate variation radiation response in different tumor regions. Total regional volume as measured by DTI at 12 months post-RT. Nine out of the 26 brain regions analyzed were significantly smaller in RT brains compared to Sham when normalized to brain volume. Data is shown as mean ± standard deviation. *p≤0.05, **p≤0.01, ***p≤0.001.
Cognitive performance

At the end of the 5-CSRTT training period, Sham animals on average completed Stage 7 which was determined to be the standard testing stage. RT animals on average completed only Stage 5. Analysis of cognitive performance at 12 months post-RT showed that there was a significant effect of RT. When tested at Stage 7 there were no significant differences between Sham and RT in percent accuracy (Sham=84±6%, RT=79±5%) and premature (Sham=8±4, RT=6±3%). RT animals however did not achieve the same performance as Shams in percent correct (Sham=51±12%, RT=34±12%, p=0.0046) and omission (Sham=32±12, RT=50±19, p=0.025) (Figure 22). This shows that RT animals are able to accurately perform the task with the same accuracy as Sham animals but did not have the same level of participation.

When tested at the level the animals achieved during training there was no significant difference between Sham and RT for percent accuracy (Sham=74±10%, RT=80±8%), percent correct (Sham=32±12%, RT=32±9%), omission (Sham=50±16, RT=59±7), or premature (Sham=9±4, RT=6±3) (Figure 21). These results seem to indicate that animals were able to remember the task and did not show further cognitive deficits. Daily breakdown of each cognitive metric can be found in Appendix C.
Figure 21. 5-CSRTT testing results at 12 months post-RT. Cognitive performance was measured at 12 months post-RT at the stage that was reached at end of training at 3 months post-RT. Percent accuracy, percent correct response, omissions and premature was averaged over the 5 day testing period. There was no significant difference between Sham and RT animals on 5-CSRTT performance. Data is shown as mean±standard deviation.
Figure 22. 5-CSRTT testing results when tested at Stage 7. Cognitive performance was measured at 12 months post-RT at Stage 7 (average stage that Sham animals reached at the end of training). Percent accuracy, percent correct response, omissions and premature was averaged over the 5 day testing period. RT animals had poor cognitive performance showing significantly lower correct response and higher omissions compared to Sham animals. Data is shown as mean±standard deviation. *p≤0.05, **p≤0.01.
Diffusion tensor imaging data

DTI data at 12 months post-RT showed a significant decrease in whole brain FA value. Significant decrease in FA value of RT brains was observed in both white matter regions such as the fimbria (Sham=0.51±0.02, RT=0.468±0.012) and corpus callosum (Sham=0.50±0.02, RT=0.45±0.03) as well as gray matter region such as the isocortex (Sham=0.266±0.012, RT=0.239±0.009. The septum was the only region that showed significant increase in RT animals compared to Sham (Sham=0.23±0.02, RT=0.254±0.012, p=0.0482) (Figure 23). This seems to indicate that there is an overall change in myelin integrity.

Changes in FA value however was not uniform across all brain regions with some regions showing more significant changes than others. FA values of individual brain regions can be found in Appendix C. Fiber number was significantly reduced in RT brains (Sham=41.3±6.3x10^3, RT=31.0±4.1 x10^3, p=0.037) while average fiber length was significantly longer (Sham=15.7±0.4, RT=18±1, p=0.0066 pixel unit (pu)) (Figure 23). There was no significant difference in the maximum fiber length (Sham=103±6, RT=98±12 pu). This seems to indicate that certain brain regions are more sensitive to radiation damage which can lead to deterioration of shorter fiber tracts leading to differences in overall fiber lengths and fiber number.
Figure 23. DTI reveals changes in brain regions and fiber values after RT. A. FA value for different regions as measured by DTI at 12 months post-RT in Sham and RT animals. RT brains had significantly lower FA value compared to Sham. Ten of the 26 brain region analyzed showed significant changes after RT. B. Whole brain fiber number value as measured using DTI at 12 months post-RT. Overall fiber number was reduced in RT animals. C. Whole brain fiber length was calculated using DTI at 12 months post-RT. Average fiber length increased after RT, pu=pixel unit. Data is shown as mean±standard deviation. Box-and-whisker plot: box extends from the 25th to 75th percentiles, median value marked by line and whiskers extend from min to max, *p≤0.05, **p≤0.01.
Connectome analysis showed changes in global measures which includes higher global efficiency (Global, p=0.035), and transitivity (Trans, p=0.033) (Figure 24). RT has also affects regional connectivity in which 38 regions were more highly connected in Sham than RT (p=0.0042).

**Figure 24. RT affects global connectomes. Error bars indicate 95% CI.**
Connectome metrics were calculated from DTI imaging at 12 months post-RT.

**White matter**

White matter (WM), gray matter (GM) and cerebral spinal fluid (CSF) segmentation using T1-weighted, T2-weighted and proton density image showed a trend towards decrease in white matter at 12 months post-RT. No significant difference between percent white matter (Sham=10.7±1.1, RT 10±3) and cerebrospinal fluid (Sham= 1.9±0.5x10^3, RT=1.3±0.51x10^3 pu) of RT and Sham
Gray matter was significantly lower in RT brains (Sham=5.0±3.0x10^3, RT=4.4±0.2x10^4 pu, p=0.00415). Reduction in gray matter matches with the reduction in volume of brain regions that are predominantly composed of gray matter.

**Figure 25. White matter analysis with standard MR images.** T1-weighted (top left), T2-weighted (top right) and proton density (bottom left) image used to segmented white matter. Final segmented map (bottom right) where green pixels are white matter, yellow is gray matter and blue is CSF.

**Imaging feature and cognitive performance correlation**

Imaging markers from DTI were highly correlated to percent accuracy at Stage 7. There were no significant correlation between cognitive performance measurement and image features on Day 1 of testing at 12 months. However, we observed a trends towards correlating average fiber length and percent accuracy. Cognitive metrics averaged over the 5 days of testing at Stage 7 showed there was a high correlation of accuracy to cluster coefficient (p=0.0153), small-world index.
(p=0.012), fiber number (p=0.0361) and FA value (p=0.0133 Figure 26). The brain's organized network may be informative of cognitive performance.

Figure 26. Correlation matrix of image features and 5-CSRTT performance. Image features derived from DTI and white matter analysis were correlated to cognitive metrics as measured at Stage 7 by 5-CSRTT, at 12 months post-RT. High correlation was found in several features including vessel number and FA value. Significant correlation (p≤0.05) is marked by an asterisk.
Inflammation

Analysis of $^{18}$FDG-PET data show significant increase in glucose uptake at 12 months post-RT (Sham SUV-mean=2.29±0.19, RT SUV-mean=3.3±0.5, p=0.00511) and was observed throughout the brain (Table 8). Analysis of the max SUV which is independent of the volume selected further confirms higher metabolic activity in RT brains (Sham SUV-max=4.3±0.5, RT SUV-max=6.4±0.7, p=0.000624) (Figure 27). Immunohistochemical analysis in the corpus callosum show that there was increase in inflammation through staining for TNFα, microglia and astrocytes in RT animals (Figure 28). Demyelination was also observed in RT brains similar to what was observed with DTI analysis (Figure 29).
Figure 27. At 12 months post-RT there was significant increase in metabolic activity in the brain. Metabolic imaging of A. Sham animal and B. RT animals with \(^{18}\)F-FDG PET at 12 months post-RT. A heat map of glucose uptake is shown where the color yellow indicates higher uptake. C. Glucose uptake was quantified using standardized uptake value (SUV) of the entire brain volume. Mean and Max SUV was higher in RT brains than Sham. Scale bar=10mm, Box-and-whisker plot: box extends from the 25th to 75th percentiles, median value marked by line and whiskers extend from min to max. *p≤0.05, ****p≤0.0001.
**Table 8.** Mean SUV of different brain regions. Data is shown as mean ± standard deviation.

<table>
<thead>
<tr>
<th>Brain Region</th>
<th>Sham</th>
<th>RT</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cerebellum</td>
<td>2.05±0.15</td>
<td>3.15±0.15</td>
<td>0.0003</td>
</tr>
<tr>
<td>Corpus Callosum</td>
<td>2.52±0.18</td>
<td>3.11±0.18</td>
<td>0.0399</td>
</tr>
<tr>
<td>Cortex</td>
<td>2.47±0.20</td>
<td>3.29±0.20</td>
<td>0.0050</td>
</tr>
<tr>
<td>Fimbria</td>
<td>2.40±0.12</td>
<td>3.3±0.5</td>
<td>0.0031</td>
</tr>
<tr>
<td>Olfactory bulb</td>
<td>2.37±0.21</td>
<td>3.5±0.5</td>
<td>0.0003</td>
</tr>
</tbody>
</table>

**Figure 28.** Neuroinflammation and demyelination is observed up to 12 months post-RT. A-B. Brain sections obtained at 12 months post-RT were stained for different markers of inflammation. Section of corpus callosum stained with myelin basic protein (MBP) shows reduction in myelin thickness after RT. C-F. An increase in TNF (tumor necrosis factor) and GFAP (glial fibrillary acidic protein) positive cells. Data is shown as mean±standard deviation. *p≤0.05, **p≤0.01.
Figure 29. Presence of microglia is significantly higher in RT brains. A. Brain sections obtained at 12 months post-RT were stained for marker of inflammation. Double stained cells are shown with a thin arrow, single stained IBA1 cells shown with large filled arrow, and single-stained TNF cells in empty arrows. B-C. Quantification of IBA-1 and double positive cells. Data is shown as mean ± standard deviation. *p≤0.05.

5.3. Discussion

In this study we have developed a preclinical model of radiation-induced cognitive deficit and using this model we identified late term imaging changes in both MRI and PET images. DTI analysis revealed regional volumetric changes and functional changes including reduction in FA value, fiber number and reorganization
of the brain's connectivity. In addition to fiber tracking, PET imaging showed increase in glucose uptake in the brain which may be driven by chronic neuroinflammation. Together these late term imaging markers provide further insight into factors that drive cognitive outcome.

The volumetric changes of the individual brain regions demonstrated that the brain is not homogenous and that some regions are more resistant than others to RT-induced damage. The volumetric changes however were not always indicative functional changes as only two brain regions, corpus callosum and striatum, showed changes in both volume and FA value. Therefore volumetric changes alone may not be informative of cognitive outcome. FA values have been established as a sensitive measures of radiation-induced brain injury, specifically in white matter changes\textsuperscript{114}. Similar to previous studies we also observed decrease FA value in whole brain and specific regions including the hippocampus\textsuperscript{113,138-140}. These changes can be correlated to cognitive performance as demonstrated by Mabbot et al. and Khong et al. in which decreased FA value was indicative of decreased cognitive performance\textsuperscript{112,141}. Using 5-CSRTT we also observed a similar trend with accuracy but in additional to FA value we also identified fiber number to be highly correlated to cognitive outcome. Even though white matter damage was observed with FA value we did not find percent white matter to be indicative of radiation-induced injury or cognitive performance in our preclinical model. Reddick et al. however found that patients treated with RT developed less normal appearing white matter
in medulloblastoma patients\textsuperscript{142}. This discrepancy may be due our choice of animal model and radiation dose.

In addition to fiber track changes we have also identified local and global changes to the brain’s network connectivity using connectome analysis. RT animals had higher transitivity, indicating increases in segregated networks and a shift towards specialized neural processing\textsuperscript{132}. Cognitive performance (accuracy) was highly was negatively correlated to abnormal network segregation which included modularity, cluster coefficient and small world index features. These features however were not statistically different between Sham and RT animals though correspondingly, accuracy was also not significantly different between the Sham and RT groups. RT seems to cause a reorganization of the brain network leading it to have more specialization of smaller sub-networks. These metrics can be used to gauge cognitive performance in an objective manner. Furthermore we’ve identified several image features that are indicative of 5-CSRTT performance which seems to suggest that there are multiple factors that can drive cognitive outcome. The identified features in this preclinical model are potential new markers for assessing patient response in clinical studies.

Neuroinflammation has been implicated as a driving force in radiation-induced brain injury. RT induces an inflammatory response with increases in TNF-\(\alpha\), microglia and astrocyte activation\textsuperscript{106,143}. This inflammatory response can persists up to months post-RT as shown in preclinical models\textsuperscript{144}. To our knowledge this is the
first report on the use of $^{18}$FDG-PET to image late-term neuroinflammation showing increase in metabolic activity. The activation of these immune cells could contribute to the signal observed. Previous studies have used this imaging technique to diagnose and identify neurodegenerative disease such as Alzheimer and Parkinson's. This would allow for longitudinal surveillance of the levels of inflammation and how these changes may affect cognitive performance. Interestingly, the increase in metabolism was observed throughout the brain, even in regions such as the cerebellum that was shielded during irradiation, which seems to indicate that neuroinflammation is systemic and not localized to exposed tissue.

There are several limitations to this study including limited sample size, timing of training period and technique used to acquire DTI data. Increasing the sample size would have helped us achieve higher statistical significance in the trend observed of decreasing white matter volume with time after irradiation. Due to the young age of the animals it was not possible to train animals prior to RT and therefore we were not able to establish a true baseline of cognitive performance. Inclusion of other cognitive tests can potentially uncover new correlation since 5-CSRTT is used to measure frontal lobe function. Despite these limitations we have identified imaging markers of late-term radiation-induced imaging changes some of which are highly correlated to cognitive performance.
5.4. Conclusions

While there are currently no biomarkers that can identify individuals susceptible to radiation-induced brain injury and the resulting impairment of their cognitive performance, medical images may provide a viable predictive tool to this end. In this study we identified several imaging markers of late-term radiation-induced damage that are highly correlated to cognitive performance. These imaging markers, derived from scans that are already acquired in the standard clinical workflow, can provide clinicians with a non-invasive tool to help predict RT treatment efficacy and the severity its side effects on a patient-to-patient basis, allowing clinicians to make more informed treatment decisions and ultimately improve treatment outcomes in radiation therapy.
Conclusion

Understanding treatment response is important especially for radiation therapy where it has a critical role in curing or managing different cancers. This is especially crucial for the pediatric patients who are most vulnerable to the lifelong side-effects of radiation therapy. The results from this work have demonstrated the potential that standard medical images can have in identifying treatment resistant tumors and therefore provide an additional source of information for clinicians to develop more personalize treatment plan, in particular when deciding the role that radiation therapy might have in the treatment.

As our study shows, medical images that are routinely acquired and are an integral part of the clinical workflow of cancer treatment, have the potential to reveal more than just anatomical information when combined with higher order imaging processing techniques. Radiomics, or the extraction of quantitative image
features with techniques such as texture analysis, when combined with machine learning techniques like the ones described in this body of work, can provide valuable new information related to disease progression and potentially redefine the role that medical images may have in the clinical workflow.

Through this work, we sought to determine the feasibility of using texture analysis to characterize tumor and normal tissue response to radiation therapy. In doing so we have established a pipeline of segmentation, extraction and construction of different classification models. Results from this work demonstrate that standard medical images are able to A) diagnose brain tumor noninvasively, B) predict treatment resistance in pediatric brain tumor model, and C) identify markers of radiation-induced brain injury.

As the utility of these medical images continue to expand and as more techniques are being developed to interpret the content of these images, there needs to be standardization of image feature extraction. We have found that the classifier performance can be greatly impacted by preprocessing techniques such as GLCM size and location where features are extracted. Furthermore the choice of classifier is also important where the same classifier may not be robust across wide ranges of conditions such as pre and post-RT. Therefore in order for these image features, as well as the models they produce, to have a significant impact on clinical practice, a great deal of consideration should be given to the method used.
While we have shown promising results in extending the use of these images there are some limitations. One of these limitations is the use of preclinical models. While we have developed a radioresistant model, further validation is needed to ensure it mimics the radioresistance that is observed in the clinic. These preclinical models however offer the ideal conditions to determine the predictive power these image features might have under the different context without confounding factors such as scanner type, protocols and tumor volume. Another limitation is the use of supervised machine learning algorithms. These algorithms rely on training data sets to construct the models, which are inherently specific to the conditions under which they were acquired. Therefore, to develop a truly robust prediction model, a larger range of training data is required unless models are tailored to specific conditions.

Future work could build on this study to extend the use of standard medical images and redefine the role of diagnostic imaging in the current clinical workflow. This study has demonstrated the feasibility of gaining new diagnostic information, using only what already acquired in the clinic, to help predict treatment outcomes. In addition to treatment response, diagnosing tumors without the need for invasive procedures can help prevent patients from being subjected to unnecessary risks associated with obtaining the diagnosis. Findings from this study as well as the work done by other researchers may potentially lead to “digital biopsy” of the tissue in which clinicians can simply use an image to make a diagnosis. This work can help complement the current clinical workflow for better treatment planning, and ultimately, improve treatment outcomes.
Appendix A

Calculation of image features:

\[ P(i, j) = \text{gray level co-occurrence matrix for given } \alpha \text{ and } \delta \]

\( N = \text{number of discrete gray intensity levels} \)

\( \mu = \text{mean of } P(i, j) \)

\( p_x(i) = \text{marginal row probabilities} \)

\( p_y(i) = \text{marginal column probabilities} \)

\( \mu_x = \text{mean of } p_x \)

\( \mu_y = \text{mean of } p_y \)

\( \sigma_x = \text{standard deviation of } p_x \)

\( \sigma_y = \text{standard deviation of } p_y \)

\[ P_{x+y}(k) = \sum_{i=1}^{N} \sum_{j=1}^{N} P(i,j), i + j = k, k = 2,3,4,...,2N \]

\[ P_{x-y}(k) = \sum_{i=1}^{N} \sum_{j=1}^{N} P(i,j), |i + j| = k, k = 0,1,2,3,...,N-1 \]

\[ H_X = - \sum_{i=1}^{N} P_x(i) \log_2[p_x(i)] \]

\[ H_Y = - \sum_{i=1}^{N} P_y(i) \log_2[p_y(i)] \]

\[ H_{XY} = - \sum_{i=1}^{N} P(i,j) \log_2[P(i,j)] \]
\[ HXY1 = - \sum_{i=1}^{N} P(i, j) \log_2[p_x(i)p_y(j)] \]

\[ HXY2 = - \sum_{i=1}^{N} [p_x(i)p_y(j)] \log_2[p_x(i)p_y(j)] \]

<table>
<thead>
<tr>
<th>Statistical Measure</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autocorrelation</td>
<td>$\sum_{i=1}^{N} \sum_{j=1}^{N} ijP(i, j)$</td>
</tr>
<tr>
<td>Cluster Prominence</td>
<td>$\sum_{i=1}^{N} \sum_{j=1}^{N} [i + j - u_x(i) - u_y(j)]^4 P(i, j)$</td>
</tr>
<tr>
<td>Cluster shade</td>
<td>$\sum_{i=1}^{N} \sum_{j=1}^{N} [i + j - u_x(i) - u_y(j)]^3 P(i, j)$</td>
</tr>
<tr>
<td>Contrast</td>
<td>$\sum_{i=1}^{N} \sum_{j=1}^{N}</td>
</tr>
<tr>
<td>Correlation (1)</td>
<td>$\frac{\sum_{i=1}^{N} \sum_{j=1}^{N} ijP(i, j) - \mu_i(i)\mu_j(j)}{\sigma_x(i)\sigma_y(j)}$</td>
</tr>
<tr>
<td>Correlation (1)</td>
<td>$\frac{\sum_{i=1}^{N} \sum_{j=1}^{N} (i - \mu_i(i))(j - \mu_j(j))P(i, j)}{\sigma_x(i)\sigma_y(j)}$</td>
</tr>
<tr>
<td>Dissimilarity</td>
<td>$\sum_{i=1}^{N} \sum_{j=1}^{N}</td>
</tr>
<tr>
<td>Energy</td>
<td>$\sum_{i=1}^{N} \sum_{j=1}^{N} P(i, j)^2$</td>
</tr>
<tr>
<td>Entropy</td>
<td>$- \sum_{i=1}^{N} \sum_{j=1}^{N} P(i, j) \log_2[P(i, j)]$</td>
</tr>
<tr>
<td>Homogeneity (1)</td>
<td>$\sum_{i=1}^{N} \sum_{j=1}^{N} \frac{P(i, j)}{1 +</td>
</tr>
<tr>
<td>Homogeneity (2)</td>
<td>[\sum_{i=1}^{N} \sum_{j=1}^{N} \frac{P(i,j)}{1 +</td>
</tr>
<tr>
<td>----------------</td>
<td>-------------------------------------------------</td>
</tr>
<tr>
<td>Maximum Probability</td>
<td>[\text{max}[P(i,j)]]</td>
</tr>
<tr>
<td>Sum of Squares</td>
<td>[\sum_{i=1}^{N} \sum_{j=1}^{N} (i - \mu)^2 P(i,j)]</td>
</tr>
<tr>
<td>Sum Average</td>
<td>[\sum_{k=2}^{2N} [kP_{x+y}(k)]]</td>
</tr>
<tr>
<td>Sum Variance</td>
<td>[\sum_{k=2}^{2N} (k - \mu_{x+y})^2 P_{x+y}(k)]</td>
</tr>
<tr>
<td>Sum Entropy</td>
<td>[-\sum_{k=2}^{2N} P_{x+y}(k) \log_2 [P_{x+y}(k)]]</td>
</tr>
<tr>
<td>Difference Variance</td>
<td>[\sum_{k=0}^{N-1} (k - \mu_{x-y})^2 [P_{x+y}(k)]]</td>
</tr>
<tr>
<td>Difference Entropy</td>
<td>[-\sum_{k=0}^{N-1} P_{x-y}(k) \log_2 P_{x+y}(k)]</td>
</tr>
<tr>
<td>Information Measure of Correlation 1</td>
<td>[\frac{H_{XY} - H_{XY1}}{\text{max}(HX, HY)}]</td>
</tr>
<tr>
<td>Information Measure of Correlation 2</td>
<td>[\sqrt{1 - e^{-2(H_{XY2} - H_{XY})}}]</td>
</tr>
<tr>
<td>Inverse Difference Normalized</td>
<td>[\sum_{i=1}^{N} \sum_{j=1}^{N} \frac{P(i,j)}{1 +</td>
</tr>
<tr>
<td>Inverse Difference Moment Normalized</td>
<td>[\sum_{i=1}^{N} \sum_{j=1}^{N} \frac{P(i,j)}{1 + (</td>
</tr>
</tbody>
</table>
## Appendix B

### Tumor region used for classification model (Validation Accuracy) – Decision Tree

<table>
<thead>
<tr>
<th># Gray levels</th>
<th>All</th>
<th>Center</th>
<th>Middle</th>
<th>Edge</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>0.686</td>
<td>0.755</td>
<td>0.741</td>
<td>0.451</td>
</tr>
<tr>
<td>16</td>
<td>0.722</td>
<td>0.776</td>
<td>0.764</td>
<td>0.457</td>
</tr>
<tr>
<td>24</td>
<td>0.727</td>
<td>0.725</td>
<td>0.73</td>
<td>0.458</td>
</tr>
<tr>
<td>32</td>
<td>0.807</td>
<td>0.767</td>
<td>0.749</td>
<td>0.559</td>
</tr>
<tr>
<td>48</td>
<td>0.812</td>
<td>0.723</td>
<td>0.731</td>
<td>0.571</td>
</tr>
<tr>
<td>64</td>
<td>0.817</td>
<td>0.712</td>
<td>0.733</td>
<td>0.505</td>
</tr>
<tr>
<td>96</td>
<td>0.819</td>
<td>0.707</td>
<td>0.745</td>
<td>0.499</td>
</tr>
<tr>
<td>128</td>
<td>0.805</td>
<td>0.671</td>
<td>0.707</td>
<td>0.55</td>
</tr>
<tr>
<td>256</td>
<td>0.819</td>
<td>0.722</td>
<td>0.677</td>
<td>0.487</td>
</tr>
<tr>
<td>512</td>
<td>0.802</td>
<td>0.744</td>
<td>0.751</td>
<td>0.461</td>
</tr>
</tbody>
</table>

### Tumor region used for classification model (Validation Accuracy) – Random Forest

<table>
<thead>
<tr>
<th># Gray levels</th>
<th>All</th>
<th>Center</th>
<th>Middle</th>
<th>Edge</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>0.789</td>
<td>0.726</td>
<td>0.647</td>
<td>0.435</td>
</tr>
<tr>
<td>16</td>
<td>0.737</td>
<td>0.753</td>
<td>0.676</td>
<td>0.450</td>
</tr>
<tr>
<td>24</td>
<td>0.776</td>
<td>0.685</td>
<td>0.735</td>
<td>0.507</td>
</tr>
<tr>
<td>32</td>
<td>0.816</td>
<td>0.740</td>
<td>0.706</td>
<td>0.435</td>
</tr>
<tr>
<td>48</td>
<td>0.803</td>
<td>0.671</td>
<td>0.676</td>
<td>0.478</td>
</tr>
<tr>
<td>64</td>
<td>0.789</td>
<td>0.712</td>
<td>0.706</td>
<td>0.450</td>
</tr>
<tr>
<td>96</td>
<td>0.776</td>
<td>0.671</td>
<td>0.706</td>
<td>0.536</td>
</tr>
<tr>
<td>128</td>
<td>0.816</td>
<td>0.753</td>
<td>0.735</td>
<td>0.493</td>
</tr>
<tr>
<td>256</td>
<td>0.842</td>
<td>0.726</td>
<td>0.765</td>
<td>0.565</td>
</tr>
<tr>
<td>512</td>
<td>0.842</td>
<td>0.740</td>
<td>0.735</td>
<td>0.522</td>
</tr>
</tbody>
</table>
### Tumor region used for classification model
(Validation Accuracy) – Support Vector Machine

<table>
<thead>
<tr>
<th># Gray levels</th>
<th>All</th>
<th>Center</th>
<th>Middle</th>
<th>Edge</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>0.825</td>
<td>0.689</td>
<td>0.734</td>
<td>0.480</td>
</tr>
<tr>
<td>16</td>
<td>0.778</td>
<td>0.718</td>
<td>0.7881</td>
<td>0.500</td>
</tr>
<tr>
<td>24</td>
<td>0.778</td>
<td>0.720</td>
<td>0.690</td>
<td>0.484</td>
</tr>
<tr>
<td>32</td>
<td>0.794</td>
<td>0.684</td>
<td>0.773</td>
<td>0.496</td>
</tr>
<tr>
<td>48</td>
<td>0.836</td>
<td>0.750</td>
<td>0.763</td>
<td>0.486</td>
</tr>
<tr>
<td>64</td>
<td>0.775</td>
<td>0.705</td>
<td>0.810</td>
<td>0.446</td>
</tr>
<tr>
<td>96</td>
<td>0.801</td>
<td>0.743</td>
<td>0.818</td>
<td>0.538</td>
</tr>
<tr>
<td>128</td>
<td>0.734</td>
<td>0.704</td>
<td>0.805</td>
<td>0.489</td>
</tr>
<tr>
<td>256</td>
<td>0.757</td>
<td>0.761</td>
<td>0.715</td>
<td>0.498</td>
</tr>
<tr>
<td>512</td>
<td>0.841</td>
<td>0.734</td>
<td>0.802</td>
<td>0.463</td>
</tr>
</tbody>
</table>
Appendix C

<table>
<thead>
<tr>
<th>Brain Region – Volume (pu)</th>
<th>Sham</th>
<th>RT</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accumbens</td>
<td>0.684±0.014</td>
<td>0.635±0.051</td>
<td>0.099</td>
</tr>
<tr>
<td>Amygdala</td>
<td>2.559±0.199</td>
<td>2.326±0.069</td>
<td>0.0271</td>
</tr>
<tr>
<td>Anterior commissure</td>
<td>0.157±0.016</td>
<td>0.156±0.015</td>
<td>0.949</td>
</tr>
<tr>
<td>Brain stem</td>
<td>0.154±0.010</td>
<td>0.149±0.007</td>
<td>0.359</td>
</tr>
<tr>
<td>Cerebellum</td>
<td>12.366±0.608</td>
<td>13.314±0.349</td>
<td>0.0133</td>
</tr>
<tr>
<td>Cingulum</td>
<td>0.208±0.009</td>
<td>0.201±0.019</td>
<td>0.480</td>
</tr>
<tr>
<td>Corpus Callosum</td>
<td>3.177±0.068</td>
<td>2.895±0.175</td>
<td>0.0165</td>
</tr>
<tr>
<td>Diagonal_domain</td>
<td>0.262±0.024</td>
<td>0.324±0.045</td>
<td>0.0355</td>
</tr>
<tr>
<td>Diencephalon</td>
<td>4.624±0.084</td>
<td>4.380±0.043</td>
<td>0.000286</td>
</tr>
<tr>
<td>Fimbria</td>
<td>0.707±0.037</td>
<td>0.712±0.036</td>
<td>0.853</td>
</tr>
<tr>
<td>Hindbrain</td>
<td>10.412±0.515</td>
<td>11.063±0.605</td>
<td>0.117</td>
</tr>
<tr>
<td>Hippocampal</td>
<td>5.330±0.055</td>
<td>5.416±0.249</td>
<td>0.523</td>
</tr>
<tr>
<td>Hypothalamus</td>
<td>2.077±0.076</td>
<td>1.900±0.105</td>
<td>0.0211</td>
</tr>
<tr>
<td>Internal_capsule</td>
<td>1.300±0.163</td>
<td>1.338±0.079</td>
<td>0.628</td>
</tr>
<tr>
<td>Isocortex</td>
<td>28.402±0.499</td>
<td>27.508±0.553</td>
<td>0.0320</td>
</tr>
<tr>
<td>Midbrain</td>
<td>4.723±0.130</td>
<td>4.527±0.293</td>
<td>0.251</td>
</tr>
<tr>
<td>Olfactory_structures</td>
<td>11.275±0.906</td>
<td>9.684±0.877</td>
<td>0.0241</td>
</tr>
<tr>
<td>Optic_pathway</td>
<td>0.305±0.037</td>
<td>0.343±0.037</td>
<td>0.150</td>
</tr>
<tr>
<td>Pallidum</td>
<td>0.763±0.020</td>
<td>0.795±0.036</td>
<td>0.147</td>
</tr>
<tr>
<td>Pineal_gland</td>
<td>0.036±0.029</td>
<td>0.040±0.007</td>
<td>0.720</td>
</tr>
<tr>
<td>Pituitary</td>
<td>0.591±0.389</td>
<td>0.209±0.040</td>
<td>0.0388</td>
</tr>
<tr>
<td>Preoptic_area</td>
<td>0.452±0.039</td>
<td>0.474±0.057</td>
<td>0.514</td>
</tr>
<tr>
<td>Septum</td>
<td>0.6345±0.016</td>
<td>0.668±0.034</td>
<td>0.109</td>
</tr>
<tr>
<td>Striatum</td>
<td>3.921±0.070</td>
<td>3.820±0.056</td>
<td>0.034</td>
</tr>
<tr>
<td>Substantia nigra</td>
<td>0.366±0.031</td>
<td>0.365±0.035</td>
<td>0.955</td>
</tr>
<tr>
<td>Ventracles</td>
<td>0.357±0.017</td>
<td>0.355±0.015</td>
<td>0.844</td>
</tr>
<tr>
<td>Brain Region</td>
<td>Sham</td>
<td>RT</td>
<td>P-value</td>
</tr>
<tr>
<td>-----------------------</td>
<td>----------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>Accumbens</td>
<td>0.201±0.011</td>
<td>0.215±0.016</td>
<td>0.188</td>
</tr>
<tr>
<td>Amygdala</td>
<td>0.273±0.025</td>
<td>0.262±0.005</td>
<td>0.291</td>
</tr>
<tr>
<td>Anterior commissure</td>
<td>0.529±0.024</td>
<td>0.507±0.009</td>
<td>0.0709</td>
</tr>
<tr>
<td>BNST</td>
<td>0.282±0.016</td>
<td>0.294±0.015</td>
<td>0.270</td>
</tr>
<tr>
<td>Cerebellum</td>
<td>0.292±0.041</td>
<td>0.253±0.019</td>
<td>0.080</td>
</tr>
<tr>
<td>Cingulum</td>
<td>0.432±0.024</td>
<td>0.407±0.015</td>
<td>0.080</td>
</tr>
<tr>
<td><strong>Corpus callosum</strong></td>
<td><strong>0.499±0.020</strong></td>
<td><strong>0.447±0.026</strong></td>
<td><strong>0.0093</strong></td>
</tr>
<tr>
<td>Diagonal_domain</td>
<td>0.230±0.014</td>
<td>0.269±0.015</td>
<td>0.331</td>
</tr>
<tr>
<td>Diencephalon</td>
<td>0.284±0.012</td>
<td>0.274±0.005</td>
<td>0.114</td>
</tr>
<tr>
<td>Fimbria</td>
<td><strong>0.505±0.022</strong></td>
<td><strong>0.468±0.012</strong></td>
<td><strong>0.0086</strong></td>
</tr>
<tr>
<td>Hindbrain</td>
<td>0.338±0.012</td>
<td>0.334±0.011</td>
<td>0.600</td>
</tr>
<tr>
<td><strong>Hippocampal</strong></td>
<td><strong>0.308±0.018</strong></td>
<td><strong>0.281±0.013</strong></td>
<td><strong>0.0225</strong></td>
</tr>
<tr>
<td>Hypothalamus</td>
<td>0.266±0.008</td>
<td>0.261±0.006</td>
<td>0.259</td>
</tr>
<tr>
<td>Internal_capsule</td>
<td>0.496±0.020</td>
<td>0.496±0.007</td>
<td>0.968</td>
</tr>
<tr>
<td>Isocortex</td>
<td><strong>0.266±0.012</strong></td>
<td><strong>0.239±0.009</strong></td>
<td><strong>0.00278</strong></td>
</tr>
<tr>
<td>Midbrain</td>
<td>0.306±0.030</td>
<td>0.278±0.012</td>
<td>0.0666</td>
</tr>
<tr>
<td><strong>Olfactory structures</strong></td>
<td><strong>0.327±0.028</strong></td>
<td><strong>0.290±0.006</strong></td>
<td><strong>0.0123</strong></td>
</tr>
<tr>
<td>Optic_pathway</td>
<td>0.571±0.018</td>
<td>0.555±0.021</td>
<td>0.243</td>
</tr>
<tr>
<td><strong>Pallidum</strong></td>
<td><strong>0.254±0.004</strong></td>
<td><strong>0.245±0.007</strong></td>
<td><strong>0.0425</strong></td>
</tr>
<tr>
<td>Pineal gland</td>
<td>0.209±0.128</td>
<td>0.180±0.034</td>
<td>0.612</td>
</tr>
<tr>
<td>Pituitary</td>
<td>0.192±0.022</td>
<td>0.183±0.020</td>
<td>0.499</td>
</tr>
<tr>
<td>Preoptic_area</td>
<td>0.232±0.023</td>
<td>0.232±0.008</td>
<td>0.967</td>
</tr>
<tr>
<td><strong>Septum</strong></td>
<td><strong>0.229±0.021</strong></td>
<td><strong>0.254±0.012</strong></td>
<td><strong>0.0482</strong></td>
</tr>
<tr>
<td>Striatum</td>
<td>0.210±0.006</td>
<td>0.191±0.009</td>
<td>0.00530</td>
</tr>
<tr>
<td>Substantia_nigra</td>
<td>0.347±0.020</td>
<td>0.358±0.016</td>
<td>0.333</td>
</tr>
<tr>
<td><strong>Ventricles</strong></td>
<td><strong>0.228±0.013</strong></td>
<td><strong>0.204±0.015</strong></td>
<td><strong>0.0332</strong></td>
</tr>
<tr>
<td><strong>Whole_Brain</strong></td>
<td><strong>0.295±0.014</strong></td>
<td><strong>0.273±0.008</strong></td>
<td><strong>0.0138</strong></td>
</tr>
</tbody>
</table>
Correlation values of image features and cognition averaged over 5 days at Stage 7.

<table>
<thead>
<tr>
<th></th>
<th>R value</th>
<th>Accuracy</th>
<th>% Correct</th>
<th>Omission</th>
<th>Premature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global</td>
<td>-0.148215</td>
<td>0.0291738</td>
<td>-0.1273931</td>
<td>0.2528142</td>
<td></td>
</tr>
<tr>
<td>Local</td>
<td>0.229998</td>
<td>0.6016045</td>
<td>-0.5196753</td>
<td>0.2528142</td>
<td></td>
</tr>
<tr>
<td>Trans</td>
<td>0.229998</td>
<td>0.6016045</td>
<td>-0.5196753</td>
<td>0.2528142</td>
<td></td>
</tr>
<tr>
<td>Mod</td>
<td>-0.6491127</td>
<td>-0.252198</td>
<td>-0.2396822</td>
<td>0.4339203</td>
<td></td>
</tr>
<tr>
<td>Cluster coeff</td>
<td>-0.7356844</td>
<td>-0.1489481</td>
<td>0.3498418</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Path length</td>
<td>0.0793359</td>
<td>0.1890009</td>
<td>-0.1128555</td>
<td>-0.2285001</td>
<td></td>
</tr>
<tr>
<td>SW index</td>
<td>-0.7315952</td>
<td>-0.1601401</td>
<td>-0.1025605</td>
<td>0.3612176</td>
<td></td>
</tr>
<tr>
<td>CSF</td>
<td>0.4261497</td>
<td>0.3936471</td>
<td>-0.2014156</td>
<td>0.2795473</td>
<td></td>
</tr>
<tr>
<td>GM</td>
<td>0.2841619</td>
<td>0.3198849</td>
<td>-0.2609261</td>
<td>0.2309338</td>
<td></td>
</tr>
<tr>
<td>%WM (ICV)</td>
<td>0.3489714</td>
<td>0.3081242</td>
<td>0.3256411</td>
<td>-0.4422767</td>
<td></td>
</tr>
<tr>
<td>Fiber #</td>
<td>0.6645839</td>
<td>0.3290049</td>
<td>-0.123831</td>
<td>-0.1094965</td>
<td></td>
</tr>
<tr>
<td>Max length</td>
<td>0.1593752</td>
<td>-0.0501182</td>
<td>0.0706764</td>
<td>-0.3156828</td>
<td></td>
</tr>
<tr>
<td>Avg. length</td>
<td>-0.2979296</td>
<td>-0.61602</td>
<td>0.4732815</td>
<td>-0.458529</td>
<td></td>
</tr>
<tr>
<td>FA</td>
<td>0.7452009</td>
<td>0.3946773</td>
<td>-0.1613925</td>
<td>-0.0905633</td>
<td></td>
</tr>
</tbody>
</table>

Correlation values of image features and cognition measurement on first day of Stage 7 testing.

<table>
<thead>
<tr>
<th></th>
<th>R value</th>
<th>Accuracy</th>
<th>% Correct</th>
<th>Omission</th>
<th>Premature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global</td>
<td>-0.2005554</td>
<td>-0.0178385</td>
<td>-0.1327695</td>
<td>0.3341282</td>
<td></td>
</tr>
<tr>
<td>Local</td>
<td>0.3722592</td>
<td>0.556555</td>
<td>-0.4494617</td>
<td>0.1967588</td>
<td></td>
</tr>
<tr>
<td>Trans</td>
<td>0.1789377</td>
<td>0.2536312</td>
<td>-0.0827745</td>
<td>-0.1766788</td>
<td></td>
</tr>
<tr>
<td>Mod</td>
<td>-0.1487562</td>
<td>-0.0494715</td>
<td>-0.1471671</td>
<td>0.3597968</td>
<td></td>
</tr>
<tr>
<td>Cluster coeff</td>
<td>-0.1016302</td>
<td>-0.1567361</td>
<td>-0.0147168</td>
<td>0.243222</td>
<td></td>
</tr>
<tr>
<td>Path length</td>
<td>0.3814396</td>
<td>0.1681164</td>
<td>-0.030802</td>
<td>-0.1822945</td>
<td></td>
</tr>
<tr>
<td>SW index</td>
<td>-0.1268371</td>
<td>-0.1663744</td>
<td>-0.0123888</td>
<td>0.2527266</td>
<td></td>
</tr>
<tr>
<td>CSF</td>
<td>0.5190361</td>
<td>0.4761841</td>
<td>-0.3241671</td>
<td>0.0672529</td>
<td></td>
</tr>
<tr>
<td>GM</td>
<td>-0.0044076</td>
<td>0.299204</td>
<td>-0.3635364</td>
<td>0.3674073</td>
<td></td>
</tr>
<tr>
<td>%WM (ICV)</td>
<td>-0.4167764</td>
<td>-0.2364831</td>
<td>0.2007575</td>
<td>-0.0931357</td>
<td></td>
</tr>
<tr>
<td>Fiber #</td>
<td>0.193257</td>
<td>0.3151515</td>
<td>-0.2061566</td>
<td>0.097253</td>
<td></td>
</tr>
<tr>
<td>Max length</td>
<td>-0.4688855</td>
<td>-0.1089841</td>
<td>0.007702</td>
<td>0.096962</td>
<td></td>
</tr>
<tr>
<td>Avg. length</td>
<td>-0.4745734</td>
<td>-0.5721895</td>
<td>0.4595234</td>
<td>-0.2151323</td>
<td></td>
</tr>
<tr>
<td>FA</td>
<td>0.2496596</td>
<td>0.3834542</td>
<td>-0.244962</td>
<td>-0.0059221</td>
<td></td>
</tr>
</tbody>
</table>
References


136 Reddick, W. E., Glass, J. O., Cook, E. N., Elkin, T. D. & Deaton, R. J. Automated segmentation and classification of multispectral magnetic resonance images of


