Radiogenomic modeling predicts survival-associated prognostic groups in glioblastoma

Nicholas Nuechterlein
10/25/2021

The University of Washington
Paul G. Allen School of Computer Science & Engineering
What is Glioblastoma?

• Most common and aggressive primary adult malignant brain tumor

• Median survival of 15 months

• Incurable because
  • Extremely heterogeneous
  • Blood brain barrier

• Last approved therapeutic agent was in 2005
What is Glioblastoma?

• Most common and aggressive primary adult malignant brain tumor

• Median survival of 15 months

• Incurable because
  • Extremely heterogeneous
  • Blood brain barrier

• Last approved therapeutic agent was in 2005
What is Glioblastoma?

• Most common and aggressive primary adult malignant brain tumor

• Median survival of 15 months

• Incurable because
  • Extremely heterogeneous
  • Blood brain barrier

• Last approved therapeutic agent was in 2005
What is Glioblastoma?

• Most common and aggressive primary adult malignant brain tumor

• Median survival of 15 months

• Incurable because
  • Extremely heterogeneous
  • Blood brain barrier

• Last approved therapeutic agent was in 2005
Patient Clinical Course

Headaches, seizures

Non-responders

Clinical Trial

Surgery + Radiation + Chemotherapy

6 weeks

Responders

Non-responders

12+ Months
Motivation for predicting short-term survivors

• Better for patients
  • Poor survivors have the most to gain from upfront trials

• Better for trials
  • Identifying poor survivors upfront can help balance clinical trial arms
  • Trials will run faster with poor survivors

• But we need to know **who** the poor survivors are **upfront**
Data

- MRI
  - Rich, global representation of tumor
  - Cheap, fast, non-invasive, repeatable
  - Volumetric
    - 255 x 255 x 155 x 4
    - (> 50 M voxels)

- Our data
  - 46 TCIA preoperative glioblastomas MRI with T1ce, FLAIR, T2, T1
Copy Number

• Captures DNA structure
• Unlike MRI: invasive, expensive, not repeatable
• 23,000 x 1 (gene-level)
• Values in \{-2, -1, 0, 1, 2\}
Poor survivor definition
Glioblastoma patients who undergo second resections live longer
Patients eligible for second resections are generally healthier, better surviving patients.
Methods

• Radiogenomics/radiomics
  • An evolving field in medical imaging that strives to equate quantitative image features with the genomic profile of pictured tissues

• Pipelines
  • Image acquisition
  • Image normalization
  • Feature extraction
  • Feature selection
  • Prediction using ML models
  • (Or end-to-end deep learning models)

• Novelty
  • Feature selection method
  • Unique clinical application
Method

Preprocessing

1. DICOM -> NIfTI
   • dcm2niix

2. Skull-strip
   • The Brain Extraction Tool (BET)

3. Co-register same-subject MRI sequences
   • FMRIB’s Linear Image Registration Tool (FLIRT) from the FMRIB Software Library (FSL)

4. Normalize/bias correct
   • N4 Bias Field Correction

Li, et al., J Neurosci Methods (2016)
Jenkinson et al. Med Image Anal (2001)
Segmentation

• U-net based architecture
• Used ESP blocks

Mehta et al., *ECCV* (2018)
Ronneberger et al., *MICCAI* (2015)
Segmentation Results

<table>
<thead>
<tr>
<th>Raw MR</th>
<th>Prediction</th>
<th>Ground Truth</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Raw MR</th>
<th>Prediction</th>
<th>Ground Truth</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Image filters / transformations

• Identity
• Laplacian of Gaussian (LoG)
• Wavelet
• Local binary patterns (LBP)
• Exponential, logarithm, square, square root

Texture analysis

• Image texture gives us information about the spatial arrangement of color or intensities in an image

• Example: Grey-level co-occurrence matrix (GLCM)
The diagonal elements all represent pixel pairs with no grey level difference.
Feature Extraction

• Histogram
  • Percentile, energy, entropy, kurtosis, skewness, uniformity, etc.

• Texture
  • GLCM (Gray Level Co-occurrence Matrix)
    • Contrast, correlation, etc.
  • GLRLM (Grey-Level Run Length Matrix)
  • GLSZM (Gray Level Size Zone Matrix)
  • GLDM (Gray Level Dependence Matrix)
  • NGTDM (Neighboring Gray Tone Difference Matrix)

• Implementation
  • pyradiomics

Putting it all together

**MR Modality**

- T1ce
- FLAIR
- T2
- T1

**Tumor compartments**

- FLAIR Abnormality
- T1ce enhancement
- Necrotic tissue

**Image Transformations**

- LoG
- Wavelets
- LBP

**Extracted Features**

- FLAIR, T2 Ab., LoG
- FLAIR, Enhancement, Wavelets
- T2, Necrotic, LBP

**Texture Features**

- Feature = \{m, c, t, e, f\}

35,000+ features per patient
Feature selection

• Feature set is far too large for modeling a few number of samples
• Feature selection overfits
  • Recursive feature elimination
  • Variance thresholding
  • LASSO feature selection

• We want to leverage the structure of our features
Feature Selection

• Stage 1
  • Aggregate a bag $B$ of LASSO-selected features, including duplicates, by training LASSO models on random subsets of the training data
Feature Selection
Feature Selection

• Stage 2 & 3
  • Use $B$ to determine which feature components ($C$) are most relevant to the classification task
  • Generate the set of 288 features whose components were determined from the set $C$
  • Use PCA to further reduce the dimensionality of our feature set to 15
Modeling

- 15 PCA Features
- Collection of small machine learning models
- Cross validation
Results

![Graph showing model inputs and ROC curve averaged over 100 trials. The graph compares models like Logistic Regression, LASSO, SVM, MLP, XGBoost, and Random Forest with different model inputs, such as Ours, LASSO FS, RFE FS, Var FS, All + PCA, All, VASARI, and Shape. The AUC values are plotted against the false positive rate in the ROC curve.](image)
Results
Discussion

• AUC > 0.80
• Attributes Not Selected
  • Enhancing tumor!
  • Identity transformation
• Attributes Selected
  • Laplacian of Gaussian transform (edge detector)
  • T2 Abnormality on FLAIR
Imaging Summary

• Developed a custom feature selection method that allows for the prediction of poor surviving glioblastoma patients, but leaves room for improvement

• Imaging limitations
  • Until scanner protocol is standardized, noise will interfere with model reliability
  • Low sample counts
  • Patients almost always get first resections, thus the fact that MRI is cheap and non-invasive is not necessarily an advantage in the upfront setting
Acknowledgments

University of Washington, Seattle, WA
- Linda Shapiro, PhD
- Beibin Li
- Mehmet Saygin Seyfioğlu
- Sachin Mehta, PhD
- Shima Nofallah

Fred Hutch Cancer Research Center, Seattle, WA
- Eric Holland, MD, PhD
- Sonali Arora

Funding Sources

University of Washington, Seattle, WA
- PJ Cimino, MD, PhD
- Michelle Casad, PhD
- James Fink, MD
- David Haynor, MD, PhD