Radiographic prediction of meningioma grade by semantic and radiomic features

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

Citation

Published Version
doi:10.1371/journal.pone.0187908

Citable link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:34493227

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA
Radiographic prediction of meningioma grade by semantic and radiomic features

Thibaud P. Coroller1*, Wenya Linda Bi2,3*, Elizabeth Huynh1, Malak Abedalthagafi4,5, Ayal A. Aizer1, Noah F. Greenwald2, Chintan Parmar1, Vivek Narayan1, Winona W. Wu2, Samuel Miranda de Moura2, Saksham Gupta2, Rameen Beroukhim3,6, Patrick Y. Wen6, Ossama Al-Mefty2, Ian F. Dunn2, Sandro Santagata4, Brian M. Alexander1, Raymond Y. Huang7‡, Hugo J. W. L. Aerts1,7‡

1 Department of Radiation Oncology, Brigham and Women’s Hospital, Dana Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts, United States of America, 2 Department of Neurosurgery, Brigham and Women’s Hospital, Harvard Medical School, Boston, Massachusetts, United States of America, 3 Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, Massachusetts, United States of America, 4 Department of Pathology Brigham and Women’s Hospital, Harvard Medical School, Boston, Massachusetts, United States of America, 5 The Saudi Human Genome Project, King Abdulaziz City for Science and Technology and Research Center at King Fahad Medical City, Riyadh, Saudi Arabia, 6 Center for Neuro-Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts, United States of America, 7 Department of Radiology, Brigham and Women’s Hospital, Harvard Medical School, Boston, Massachusetts, United States of America

*These authors contributed equally to this work.
‡ These authors also contributed equally to this work.

Abstract

Objectives

The clinical management of meningioma is guided by tumor grade and biological behavior. Currently, the assessment of tumor grade follows surgical resection and histopathologic review. Reliable techniques for pre-operative determination of tumor grade may enhance clinical decision-making.

Methods

A total of 175 meningioma patients (103 low-grade and 72 high-grade) with pre-operative contrast-enhanced T1-MRI were included. Fifteen radiomic (quantitative) and 10 semantic (qualitative) features were applied to quantify the imaging phenotype. Area under the curve (AUC) and odd ratios (OR) were computed with multiple-hypothesis correction. Random-forest classifiers were developed and validated on an independent dataset (n = 44).

Results

Twelve radiographic features (eight radiomic and four semantic) were significantly associated with meningioma grade. High-grade tumors exhibited necrosis/hemorrhage (OR$_{\text{sem}}$ = 6.6, AUC$_{\text{rad}}$ = 0.62–0.68), intratumoral heterogeneity (OR$_{\text{sem}}$ = 7.9, AUC$_{\text{rad}}$ = 0.65), nonspherical shape (AUC$_{\text{rad}}$ = 0.61), and larger volumes (AUC$_{\text{rad}}$ = 0.69) compared to low-grade tumors. Radiomic and semantic classifiers could significantly predict meningioma grade (AUC$_{\text{sem}}$ = 0.76 and AUC$_{\text{rad}}$ = 0.78). Furthermore, combining them increased the
classification power ($\text{AUC}_{\text{radio}} = 0.86$). Clinical variables alone did not effectively predict tumor grade ($\text{AUC}_{\text{clin}} = 0.65$) or show complementary value with imaging data ($\text{AUC}_{\text{comb}} = 0.84$).

**Conclusions**

We found a strong association between imaging features of meningioma and histopathologic grade, with ready application to clinical management. Combining qualitative and quantitative radiographic features significantly improved classification power.

**Introduction**

Meningiomas are the most common primary brain tumor in adults, with most considered benign by the World Health Organization histopathologic criteria (WHO grade I)[1,2]. A distinct and increasing proportion of meningiomas are deemed high-grade (WHO grade II-III) and recur despite aggressive treatment, leading to substantial morbidity. Standard-of-care management typically involves surgical resection and often radiation therapy for high-grade (grade II-III) or progressive tumors.

Currently, the assessment of tumor grade occurs once a mass is resected and histopathological review is performed. Upon detection of a mass lesion that displays radiological features suggestive of meningioma, reliable parameters do not exist that can predict tumor grade and the associated clinical course. For example, clinical information such as age and gender show poor association with grade. Non-invasive and early predictors of meningioma grade may enhance clinical decision-making by providing prognostic information that could guide the decision of whether to observe or to treat.

The radiographic appearance of a tumor can be described using both quantitative and qualitative measures (Fig 1). Radiomics is an emerging field of quantitative imaging focused on leveraging large sets of imaging features to create an atlas[3–9] that would foster the automatic, reproducible, and unbiased assessment of active clinical cases[10–13]. In comparison, semantic features are tumor traits (e.g. bone invasion, necrosis) that are assessed visually by radiologists. While semantic features are highly intuitive, they are inherently subject to inter-observer variability. Both radiomic[14–29] and semantic[30–33] features have been applied as prognostic biological signatures, and therefore, may offer complementary streams to predict clinical status.

In this study, we investigated the value of radiomic and semantic imaging features for predicting the histologic grade of meningiomas from preoperative gadolinium-enhanced T1-weighted MRI.

**Methods**

This study was reviewed and approved by the Brigham and Women’s Hospital institutional review boards (IRB). Patient consent was waived by IRB protocol. All methods were performed in accordance with the relevant guidelines and regulations.

**Patient data**

A total of 181 meningiomas resected at our institution between 2003 and 2014 were reviewed for histopathology and imaging. Pre-operative gadolinium-enhanced T1-weighted MRI
sequences were chosen for analysis to represent the most frequently reviewed images for meningiomas. Six cases with motion artifact were excluded from analyses.

Histopathologic review of all tumors was performed by two board-certified neuropathologists (S.S., M.A.). Meningiomas were graded according to the 2007 World Health Organization (WHO) classification system [1]. The data were additionally reviewed according to the 2016 WHO classification system to assess any potential impact that the inclusion of brain invasion as a formal diagnostic criterion for grade II meningiomas might have on their association with imaging features. In this study, low and high grade refers to grade I and grade II/III, respectively. Atypical features for meningiomas were individually tabulated [34].

Image-based phenotyping

In this study, semantic (qualitative) and radiomic (quantitative) feature quantification was applied to preoperative MRI (Fig 1B, Table 1). The standard preoperative imaging protocol for intracranial tumors include a high-resolution gadolinium-enhanced T1-weighted 3D MPRAGE or SPGR sequence, acquired on a 1.5T or 3T scanner. For patients who had had serial imaging prior to surgery, we analyzed the MRI that was acquired closest to the date of surgery. We exported images into 3D Slicer[35] for editing and reconstructed meningioma volumes from the manual contours of individual axial MRI slices performed by two fully trained neurosurgeons familiar with the radiographic appearance of the meningiomas to limit inclusion of radiologic artifacts. All contours were reviewed by an experienced neuroradiologist (R.H.) to ensure standardization of contouring criteria across the dataset. We applied
image processing prior to feature extraction to reduce noise (mean +/- 3 standard deviations) according to well-established MRI-normalization methods. We resampled the voxel dimensions using 3x3x3 mm$^3$ as the common spacing.

Semantic features such as speculation and mass effect are MRI characteristics regularly assessed during the standard evaluation of images from patients with meningiomas. Ten binary (to avoid high inter-observer variability) semantic features were scored by an experienced neuroradiologist (R.H.) whereas radiomic features were extracted from images using a custom Matlab script. A total of 1,055 radiomic features were computed that quantify the tumor phenotype (description in S1 File). We selected fifteen features for this study based on their variance and correlation (Fig 2).

Additionally, we included two tumor size features (maximum axial diameter and volume) in the clinical data set, along with age, gender, and radiation induced status. Tumor location was also classified by two skull base surgery trained neurosurgeons, based on the origin of the

### Table 1. Description of radiographic features and filters

<table>
<thead>
<tr>
<th>Type</th>
<th>Group</th>
<th>Feature / Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radiographic features</td>
<td>Semantic</td>
<td>Intratumoral heterogeneity</td>
<td>Heterogeneity in hyperintensity of MRI signal throughout tumor</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Multilocality</td>
<td>Non-contiguous growth of tumor</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Midline shift</td>
<td>Shift of the brain past midline</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sinus invasion</td>
<td>Presence of venous sinus invasion</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Necrosis / Hemorrhage</td>
<td>Presence of necrosis or hemorrhage</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mass effect</td>
<td>Shift in normal brain parenchyma due to tumor</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cystic component</td>
<td>Fluid filled cysts within the tumor</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bone invasion</td>
<td>Appearance of tumor invading the skull</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hyperostosis</td>
<td>Bony overgrowth adjacent to tumor</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Spiculation</td>
<td>Irregularities in tumor shape and border</td>
</tr>
<tr>
<td></td>
<td>Radiomic</td>
<td>Median</td>
<td>Median voxel intensity value</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>Mean voxel intensity value</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Minimum</td>
<td>Minimal voxel intensity value</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Skewness</td>
<td>Describes the shape of a probability distribution of the voxel intensity histogram</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Spherical Disproportion (SD)</td>
<td>How different is the tumor is to a sphere with a similar volume</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cluster Prominence (CP)</td>
<td>Sensitive to flat zones (area of similar intensity)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Difference Entropy (DE)</td>
<td>Complexity of the pattern (high entropy for high number of unique patterns)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inverse Difference Normalized (IDN)</td>
<td>Sensitive to homogeneity in the tumor</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Run Length Non-uniformity (RLN)</td>
<td>Measure of heterogeneity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Short Run Low Gray-Level Emphasis (SRLGLE)</td>
<td>Measure of heterogeneity sensitive to low intensity pattern</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High Intensity Large Area Emphasis (HILAE)</td>
<td>Sensitive to flat zones with high intensity voxels (e.g. areas of hemorrhage)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Low Intensity Large Area Emphasis (LILAE)</td>
<td>Sensitive to flat zones with low intensity voxels (e.g. areas of necrosis)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Low Intensity Small Area Emphasis (LISAE)</td>
<td>Sensitive to small flat zones with low intensity voxels</td>
</tr>
<tr>
<td>Filters</td>
<td>Wavelet</td>
<td>High (L), Low (L)</td>
<td>Wavelet filters decompose images by high (increase details) and low (smooth image, leaving general shape) for every spatial component (x,y,z)</td>
</tr>
<tr>
<td></td>
<td>LoG</td>
<td>Sigma (σ)</td>
<td>Laplacian of Gaussian is a filter that highlights textures using a variable size radius (σ). Depending on the radius (from 0.5mm to 5mm with 0.5 increment), it emphasizes image textures from fine to coarse.</td>
</tr>
</tbody>
</table>

[https://doi.org/10.1371/journal.pone.0187908.t001](https://doi.org/10.1371/journal.pone.0187908.t001)
meningioma, and clustered into five groups for purposes of analysis: 1) midline skull base, 2) lateral skull base, 3) midline convexity, 4) lateral convexity or 5) other.

Univariate analysis

All statistical analyses were performed in R software version 3.3.1[36]. Our primary endpoint was the potential applicability of radiographic features to predict meningioma grade. The predictive power of semantic features (binary) was evaluated using the odds ratio (OR) and Fisher’s exact test. The predictive power of radiomic features (continuous) was analyzed using the area under the receiver operator characteristics curve (AUC) using the “survcomp” package[37] and Noether’s test.

Additionally, prediction of low grade (grade I) with the presence of atypical features was studied. A subset of the cohort with only low grade meningioma was analyzed, where we compared patients with one or more atypical features (including spontaneous necrosis, high nuclear-to-cytoplasmic ratio, prominent nucleoli, and sheet-like growth) versus none of these features using the same imaging features as for grade prediction. Hypercellularity was almost ubiquitously observed across the meningioma cohort, and therefore, not included as an atypical feature for purposes of analysis.

Finally, the association between radiomic and semantic features was investigated using the AUC. Every semantic feature was predicted by each of the radiomic features in a univariate manner. All p-values were adjusted for multiple hypothesis testing using the false discovery rate method[38].

Multivariate analysis

A temporal split was used to assign patients to a training or validation dataset. Feature selection was based on the training dataset, to ensure independence from the validation dataset (Table A in S1 File). Differences in clinical variables between datasets were assessed using the Fisher’s exact test (for categorical variables) and the Wilcoxon test (for continuous variables).

We investigated five models for grade classification based on: 1) clinical, 2) location, 3) semantic, 4) radiomic, 5) radiographic (combined radiomic and semantic features), and 6) a combined model integrating all features above. Classifications were made using the random forest method from the “randomForest” package [39]. Nested cross validation was used for model tuning and training using the “caret” package on the training set [40], leaving the validation dataset independent from the model selection process. Differences in predictive power between models were assessed using bootstrapping (1,000 iterations).
Results

Clinical cohort

Our cohort of 175 patients was mainly composed of female patients (62%), with a median age of 57 years (Table 2). 59% of cases were low-grade and 41% were high-grade. No differences were observed in WHO grade (p = 0.48), radiation-induced status (p = 0.51), or gender (p = 0.15) between the training and validation datasets.

Radiographic associations with meningioma grade or atypical features

First, examination of individual semantic (qualitative) features revealed significant associations between meningioma grade and four features (Fig 3A.1, Table 3). These features included intratumoral heterogeneity (OR = 7.9, p < 0.001), necrosis/hemorrhage (OR = 6.6, p = 0.01), venous sinus invasion (OR = 2.9, p = 0.02), and mass effect (OR = 2.3, p = 0.042). Interestingly, cystic component was not significantly associated with grade despite a high OR (6.8, p = 0.13), which is likely due a low incidence of events (6 cases, 3.4%), which introduces a high margin error. All significant features had an OR greater than one, indicating that higher grade corresponds to an increased incidence of the feature.

Second, we investigated the relationship between radiomic (quantitative) features and meningioma pathology (Fig 3A.2, Table 4). Eight radiomic features were significant from random in their association with tumor grade (range AUC = 0.59 to 0.65, p < 0.05). The best performing radiomic feature, high intensity large area emphasis (HILAE), was associated with high grade meningioma (AUC = 0.69, p < 0.001). HILAE is sensitive to large zones with high intensities (e.g. hemorrhage). In addition, low intensity large area emphasis (LILAE) was also associated with high grade meningioma (AUC = 0.63, p = 0.008) and is sensitive to large areas of low intensities (e.g. necrosis). These suggest that hemorrhagic or necrotic tumors were more likely to be high grade, consistent with the semantic feature analysis. High values of spherical disproportion (SD), which measures the degree of deviation of a tumor’s shape from...
Fig 3. A) Heatmap of the predictive power of (1) semantic and (2) radiomic features for meningioma grade (n = 175) or presence of histopathologic atypia in low grade meningiomas (n = 103). B) The association between semantic and radiomic features was investigated. Every semantic feature was predicted with each of the radiomic feature in a univariate manner that indicates their relationship. * indicates significance from random after multiple correction.

https://doi.org/10.1371/journal.pone.0187908.g003
a sphere of similar volume, and run length non-uniformity (RLN), which is sensitive to heterogeneity, were both significantly associated with high-grade tumors (AUC = 0.61, p = 0.012 and AUC = 0.65, p = 0.002, respectively).

Additionally, we examined the ability of imaging to distinguish low grade meningiomas with (n = 69) and without (n = 34) one of four atypical features (Fig 3A, Tables B-C in S2 File). While intratumoral heterogeneity and multifocality carried an OR of 1.7 and 2.1, respectively, no significant association was observed between semantic features and the presence of atypical features. In comparison, four radiomic features were significantly associated with atypical features. These features included voxel mean intensity (AUC = 0.68), low intensity small area emphasis (LISAE) (AUC = 0.66), difference entropy (DE) (AUC = 0.66), and cluster prominence (CP) (AUC = 0.65). LISAE indicated that hypointense tumors were more likely to be low grade with atypical features. None of the low grade meningiomas had a cystic component; therefore, this semantic feature was not investigated in this analysis.

Table 3. Univariate results for the semantic features. Odds ratio, lower and higher 95% confidence interval and p-value (with multiple testing correction) are reported for each feature.

<table>
<thead>
<tr>
<th>Features</th>
<th>Odds Ratio</th>
<th>95% Conf. Int.</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyperostosis</td>
<td>0.35</td>
<td>0.08</td>
<td>1.16</td>
</tr>
<tr>
<td>Spiculation</td>
<td>0.47</td>
<td>0.01</td>
<td>6.01</td>
</tr>
<tr>
<td>Multifocality</td>
<td>0.89</td>
<td>0.22</td>
<td>3.23</td>
</tr>
<tr>
<td>Bone Invasion</td>
<td>1.00</td>
<td>0.31</td>
<td>3.09</td>
</tr>
<tr>
<td>Midline Shift</td>
<td>1.39</td>
<td>0.71</td>
<td>2.70</td>
</tr>
<tr>
<td>Mass Effect</td>
<td>2.31</td>
<td>1.20</td>
<td>4.53</td>
</tr>
<tr>
<td>Sinus Invasion</td>
<td>2.91</td>
<td>1.33</td>
<td>6.59</td>
</tr>
<tr>
<td>Necrosis / Hemorrhage</td>
<td>6.60</td>
<td>1.69</td>
<td>37.88</td>
</tr>
<tr>
<td>Cystic</td>
<td>7.53</td>
<td>0.82</td>
<td>362.79</td>
</tr>
<tr>
<td>Intratumoral Heterogeneity</td>
<td>7.95</td>
<td>3.62</td>
<td>18.83</td>
</tr>
</tbody>
</table>

https://doi.org/10.1371/journal.pone.0187908.t003

Table 4. Univariate results for the radiomic features. AUC, lower and higher 95% confidence interval and p-value (with multiple testing correction) are reported for each feature.

<table>
<thead>
<tr>
<th>Features</th>
<th>AUC</th>
<th>95% Conf. Int.</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HHL Skewness</td>
<td>0.51</td>
<td>0.40</td>
<td>0.57</td>
</tr>
<tr>
<td>HLM Median</td>
<td>0.52</td>
<td>0.44</td>
<td>0.61</td>
</tr>
<tr>
<td>LoG5 Low Intensity Small Area Emp</td>
<td>0.54</td>
<td>0.46</td>
<td>0.63</td>
</tr>
<tr>
<td>LLH Short Run Low Gray Level Emp</td>
<td>0.55</td>
<td>0.37</td>
<td>0.54</td>
</tr>
<tr>
<td>Difference Entropy</td>
<td>0.56</td>
<td>0.35</td>
<td>0.52</td>
</tr>
<tr>
<td>HHH Mean</td>
<td>0.58</td>
<td>0.50</td>
<td>0.66</td>
</tr>
<tr>
<td>LoG4 High Intensity Large Area Emp</td>
<td>0.59</td>
<td>0.50</td>
<td>0.67</td>
</tr>
<tr>
<td>HLL Cluster Prominence</td>
<td>0.60</td>
<td>0.51</td>
<td>0.68</td>
</tr>
<tr>
<td>Spherical disproportion</td>
<td>0.61</td>
<td>0.53</td>
<td>0.69</td>
</tr>
<tr>
<td>LoG5 Inv. Diff. Normalized</td>
<td>0.61</td>
<td>0.53</td>
<td>0.70</td>
</tr>
<tr>
<td>HLH Low Intensity Large Area Emp</td>
<td>0.63</td>
<td>0.54</td>
<td>0.71</td>
</tr>
<tr>
<td>HHL Mean</td>
<td>0.63</td>
<td>0.55</td>
<td>0.71</td>
</tr>
<tr>
<td>Run Length Non-uniformity</td>
<td>0.65</td>
<td>0.56</td>
<td>0.73</td>
</tr>
<tr>
<td>Minimum</td>
<td>0.65</td>
<td>0.57</td>
<td>0.73</td>
</tr>
<tr>
<td>HHH High Intensity Large Area Emp</td>
<td>0.69</td>
<td>0.61</td>
<td>0.77</td>
</tr>
</tbody>
</table>

https://doi.org/10.1371/journal.pone.0187908.t004
We investigated the link between radiomic and semantic features. We found a median AUC of 0.57 (range: 0.50–0.89) between these two categories of features (Fig 3B, Table D in S2 File). A significant interaction between particular pairs of features was found (p < 0.05). Spherical disproportion (SD) was associated with mass effect (AUC = 0.61), spiculation (AUC = 0.89), and invasion of bone and venous sinus (AUC of 0.74 and 0.69, respectively). Textural features, such as high intensity large area emphasis (HILAE), run length non-uniformity (RLN) and short run length gray-level (SRLGL) were associated with tumor heterogeneity (AUC = 0.65–0.72), cystic component (AUC = 0.71–0.84), and hemorrhage / necrosis (AUC = 0.70–0.76).

Given that radiomic and semantic analyses each provide a distinct quantification of the tumor phenotype, we explored whether combining radiomic and semantic features may be synergistic in predicting meningioma grade (Fig 4, Table 5).

### Table 5. Meningioma classification validation (n = 44) for each model is reported. AUC, lower and higher 95% confidence interval and p-value (from random) are reported for each features.

<table>
<thead>
<tr>
<th></th>
<th>AUC</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>95% Conf. Int.</th>
<th>p-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical</td>
<td>0.651786</td>
<td>0.928571</td>
<td>0.375</td>
<td>0.468065, 0.835507</td>
<td>0.105389</td>
</tr>
<tr>
<td>Location</td>
<td>0.669643</td>
<td>0.892857</td>
<td>0.3125</td>
<td>0.541902, 0.907802</td>
<td>0.016002</td>
</tr>
<tr>
<td>Semantic</td>
<td>0.788973</td>
<td>0.75</td>
<td>0.625</td>
<td>0.648494, 0.913278</td>
<td>3.21E-05</td>
</tr>
<tr>
<td>Radiomic</td>
<td>0.779018</td>
<td>0.928571</td>
<td>0.625</td>
<td>0.639093, 0.918943</td>
<td>9.30E-05</td>
</tr>
<tr>
<td>Radiographic (Loc. + Sem. + Rad.)</td>
<td>0.860491</td>
<td>0.821429</td>
<td>0.625</td>
<td>0.762583, 0.960012</td>
<td>7.31E-13</td>
</tr>
<tr>
<td>Combined (All)</td>
<td>0.83817</td>
<td>0.892857</td>
<td>0.375</td>
<td>0.733099, 0.944753</td>
<td>3.45E-10</td>
</tr>
</tbody>
</table>

https://doi.org/10.1371/journal.pone.0187908.t005
A model based on clinical data, composed of information available to a clinician prior to MR imaging, did not associate with meningioma grade (AUC = 0.65, p = 0.11). In comparison, location (AUC = 0.67), semantic (AUC = 0.77) and radiomic (AUC = 0.79) models independently classified meningioma grade (p < 0.05). Interestingly, a radiographic model that combined both radiomic and semantic features showed an increased performance in the classification of tumor grade (AUC = 0.86, p < 0.001). Despite the fact that the performance of the radiographic model was higher than the semantic and radiomic models, it was not significantly better than each alone (p-value = 0.23–0.32). Lastly, adding clinical and location data to the radiographic model did not improve the performance (AUC = 0.84) compared to radiographic features (AUC = 0.86).

Additionally, we verified the validity of the imaging association with meningioma grade using the 2016 WHO guidelines, which includes brain invasion as a formal inclusion criterion for grade II. We observed similar results with the updated WHO criteria as the 2007 criteria, with additional observation of a significant association between the clinical model and pathologic grade (in Fig A, Table E in S2 File), attesting to the robustness of the radiographic association with tumor grade.

Discussion

Meningioma grade is a powerful predictor of clinical outcome and therefore influences patient management, including the decision of whether to observe, operate, or administer adjuvant therapies. Currently, tumor grade can only be determined following surgical resection and histopathological review[41]. A better approach would allow clinicians to discriminate low and high grade meningiomas before surgery, thereby facilitating management decisions and counseling at an earlier stage of clinical care. Such a shift in the diagnostic paradigm would have substantial implications for patient management, particularly in the increasingly common scenario era in which asymptomatic meningiomas are incidentally diagnosed on imaging performed for unrelated reasons. In our study, we sought to develop and test methodologies for the pre-operative diagnostic assessment of meningioma grade using two categories of radiographic data (semantic and radiomic) derived from T1-weighted contrast-enhanced MRI.

We observed strong associations between specific radiographic features and meningioma histologic grade. In particular, heterogeneous tumors with necrosis and/or hemorrhage, and irregularly shaped (non-spherical) tumors were more likely to be higher grade on univariate analysis. Two radiomic features, HILAE and LILAE, were sensitive to high and low intensity large areas, respectively, which are commonly indicative of hemorrhage and necrosis on MR images. Interestingly, both semantic and radiomic features were significantly associated with these traits and their presence indicated an increased likelihood of a high grade tumor. Tumor heterogeneity was also significantly associated with more aggressive meningioma grade in both semantic and radiomic feature analyses.

Irregularities in the shape of meningiomas such as “mushrooming” has been previously associated with high grade tumor in multiple studies[42–45]. Meningioma heterogeneity, on the other hand, is a more complex tumor trait that may be accounted for by a variety of underlying causes, including intratumoral necrosis, cystic degeneration, heterogeneous tumor cell expansion, variability in cell density, and hemorrhage[46,47]. Tumor radiographic heterogeneity has been extensively studied in glioblastoma, lung cancer, renal cell cancer, and other systemic malignancies and is felt to contribute significantly to treatment resistance and disease relapse[48]. Awareness of tumor heterogeneity may play an important role in assessing treatment response in meningiomas as well, given recent and impending clinical trials assessing novel targeted and immune therapies for aggressive meningiomas[49–51].
We further confirmed the ability of radiographic features to classify meningioma grade on an independent validation dataset. Moreover, even though semantic and radiomic features capture some common traits in the tumor (e.g. heterogeneity), the information contained in these “common” features were complementary. Merging both feature sets significantly improved classification performance, indicating an additive effect between qualitative and quantitative imaging analyses. Additionally, tumor location alone was able to classify tumor grade but was not able to improve the combined model once added to the radiographic and clinical information (likely mostly driven by the radiographic features). Clinical data, added for comparison, did not classify patients well nor did it add power to the radiographic model in the validation, although this could be limited by selection bias in the variables analyzed.

Associations between tumor characteristics and pre-operative images have been previously investigated[52]. However, no sets of phenotypic features have been consistently demonstrated to significantly associate with meningioma grade across studies. Differences between meningioma and low grade glioma was investigated using imaging features from T1-weighed and DWI[53], however, the study presented several limitations including a small sample size (n = 15). Some studies investigating imaging features suggest that benign tumors display higher ADC while malignant tumors have lower ADC values[54–56], while others fail to corroborate a similar relationship[57–59]. These conflicting results may be due to technical factors, such as the region of interest (ROI) defined and feature standardization[59,60].

Likewise, our study faces several limitations. Variations in image acquisition and quality can influence quantitative analyses. We attempted to standardize the uniformity of scans by resampling all images with a common voxel spacing to ensure dimension homogeneity and by filtering voxel intensities to reduce outlier values. Additionally, the semantic features are reported by human thus are subject to inter-observer variability. To reduce this variation, we used largely binary features to simply the output, as compared to a more complex scale (such as a score from 1 to 5) which have been suggested to associate with more inter-observer variability [61]. We used a temporal split to obtain an independent validation dataset, with comparable demographics between the cohorts, in attempt to internally validate our results. External validation from multiple institutions would strengthen these observations in the future. Our clinical model was predicated on common non-radiographic variables that may influence tumor behavior, but may reflect selection bias and data availability in this single-institution cohort.

In conclusion, we found a radiographic signature for meningioma grade using standard pre-operative contrast-enhanced MR images. We demonstrated that there is a strong link between the radiographic phenotype of a tumor and its pathology, which may provide a useful tool for precision medicine. Early and accurate prediction of meningioma grade may influence the decision to observe a tumor or to pursue surgery and earlier consideration of adjuvant therapies. Our study highlights the potential clinical impact of integrative imaging analysis in guiding meningioma management.

Supporting information

S1 File. Table A. Description of the training radiomic set. Individual description is given for every features.

S2 File. Table B Univariate results for the semantic features. Odds ratio, lower and higher 95% confidence interval and p-value (with multiple testing correction) are reported for each features. Table C Univariate results for the radiomic features. AUC, lower and higher 95%
confidence interval and p-value (with multiple testing correction) are reported for each feature. **Table D.** Association between radiomic and semantic features was investigated using AUC. **Table E.** Meningioma classification validation (n = 44) for each model is reported using the WHO 2016. AUC, lower and higher 95% confidence interval and p-value (from random) are reported for each features. **Fig A.** Area under the curve (AUC) from random forest models on the independent validation set (n = 44) for meningioma WHO 2016 grade classification. “*” indicates p-value <0.05, “**” indicates p-value <0.0001 from random prediction (Noether test).

(DOCX)

**Author Contributions**

**Conceptualization:** Thibaud P. Coroller, Wenya Linda Bi, Elizabeth Huynh, Ayal A. Aizer, Noah F. Greenwald, Chintan Parmar, Rameen Beroukhim, Patrick Y. Wen, Ossama Al-Mefty, Ian F. Dunn, Sandro Santagata, Brian M. Alexander, Raymond Y. Huang, Hugo J. W. L. Aerts.

**Data curation:** Thibaud P. Coroller, Wenya Linda Bi, Malak Abedalthagafi, Ayal A. Aizer, Noah F. Greenwald, Vivek Narayan, Winona W. Wu, Samuel Miranda de Moura, Saksham Gupta, Rameen Beroukhim, Ian F. Dunn, Sandro Santagata, Raymond Y. Huang.

**Formal analysis:** Thibaud P. Coroller, Noah F. Greenwald, Chintan Parmar.

**Funding acquisition:** Patrick Y. Wen, Raymond Y. Huang, Hugo J. W. L. Aerts.

**Investigation:** Thibaud P. Coroller, Wenya Linda Bi, Elizabeth Huynh, Malak Abedalthagafi.


**Project administration:** Patrick Y. Wen, Raymond Y. Huang.


**Software:** Vivek Narayan.

**Supervision:** Thibaud P. Coroller, Wenya Linda Bi, Elizabeth Huynh, Ayal A. Aizer, Noah F. Greenwald, Chintan Parmar, Rameen Beroukhim, Patrick Y. Wen, Ossama Al-Mefty, Ian F. Dunn, Sandro Santagata, Brian M. Alexander, Raymond Y. Huang, Hugo J. W. L. Aerts.

**Validation:** Thibaud P. Coroller, Wenya Linda Bi, Chintan Parmar.

**Visualization:** Thibaud P. Coroller, Noah F. Greenwald, Chintan Parmar, Vivek Narayan.

**Writing – original draft:** Thibaud P. Coroller, Wenya Linda Bi, Elizabeth Huynh, Malak Abedalthagafi, Ayal A. Aizer, Noah F. Greenwald, Chintan Parmar, Vivek Narayan, Winona W. Wu, Samuel Miranda de Moura, Saksham Gupta, Rameen Beroukhim, Patrick Y. Wen, Ossama Al-Mefty, Ian F. Dunn, Sandro Santagata, Brian M. Alexander, Raymond Y. Huang, Hugo J. W. L. Aerts.

**Writing – review & editing:** Thibaud P. Coroller, Wenya Linda Bi, Raymond Y. Huang, Hugo J. W. L. Aerts.
References


Radiographic prediction of meningioma grade

LABORATORY INVESTIGATION. NATURE PUBLISHING GROUP 75 VARICK ST, 9TH FLR, NEW YORK, NY 10013–1917 USA; 2014. p. 433A–433A.


