Neuro-Oncology
XX(XX), 1–9, 2020 | doi:10.1093/neuonc/noaa007 | Advance Access date 20 January 2020

Use of radiomics for the prediction of local control of brain metastases after stereotactic radiosurgery

Andrei Mouraviev, Jay Detsky, Arjun Sahgal, Mark Ruschin, Young K. Lee, Irene Karam, Chris Heyn, Greg J. Stanisz, and Anne L. Martel

Department of Medical Biophysics, University of Toronto (A.M., G.J.S., A.L.M.), Physical Sciences, Sunnybrook Research Institute (G.J.S., A.L.M.), and Odette Cancer Center, Sunnybrook Health Sciences Centre (J.D., A.S., M.R., Y.K.L., I.K., C.H.), Toronto, Ontario, Canada

Corresponding Author: Andrei Mouraviev, 2075 Bayview Ave., M6 609, Toronto, ON, Canada, M4N 3M5 (martal.lab.neuro.onc@gmail.com).

Abstract

Background. Local response prediction for brain metastases (BM) after stereotactic radiosurgery (SRS) is challenging, particularly for smaller BM, as existing criteria are based solely on unidimensional measurements. This investigation sought to determine whether radiomic features provide additional value to routinely available clinical and dosimetric variables to predict local recurrence following SRS.

Methods. Analyzed were 408 BM in 87 patients treated with SRS. A total of 440 radiomic features were extracted from the tumor core and the peritumoral regions, using the baseline pretreatment volumetric post-contrast T1 (T1c) and volumetric T2 fluid-attenuated inversion recovery (FLAIR) MRI sequences. Local tumor progression was determined based on Response Assessment in Neuro-Oncology–BM criteria, with a maximum axial diameter growth of >20% on the follow-up T1c indicating local failure. The top radiomic features were determined based on resampled random forest (RF) feature importance. An RF classifier was trained using each set of features and evaluated using the area under the receiver operating characteristic curve (AUC).

Results. The addition of any one of the top 10 radiomic features to the set of clinical features resulted in a statistically significant ($P < 0.001$) increase in the AUC. An optimized combination of radiomic and clinical features resulted in a 19% higher resampled AUC (mean = 0.793; 95% CI = 0.792–0.795) than clinical features alone (0.669, 0.668–0.671).

Conclusions. The increase in AUC of the RF classifier, after incorporating radiomic features, suggests that quantitative characterization of tumor appearance on pretreatment T1c and FLAIR adds value to known clinical and dosimetric variables for predicting local failure.

Key Points

1. Radiomic features aid local control prediction of brain metastases treated with radiosurgery.
2. Radiomic features are complementary to clinical features for this task.
3. Tumor core sphericity appears to be an important feature.

Brain metastases (BM) are common in patients with metastatic cancer and can be life-threatening. MRI is the gold standard modality for the diagnosis and follow-up of BM. Volumetric contrast-enhanced T1-weighted MRI (T1c) is used for radiotherapy treatment planning as BM appear hyperintense on T1c images due to disruptions in the tumor vasculature allowing for gadolinium uptake, while necrotic regions within the tumor appear hypointense.

Importance of the Study

Predicting whether BM will respond to SRS is a major challenge in neuro-oncology, and development of novel methods such as the integration of radiomics has potential to improve patient management as identification of BM at risk for local failure may warrant surgery, a higher radiation dose, and/or change in systemic therapeutic management. This is of particular importance for BM <1.0 cm in maximal diameter as existing criteria for predicting response are not defined. Our machine learning model utilizes radiomic features, extracted from the baseline volumetric T1c and volumetric T2 fluid-attenuated MR images and improves the prediction of local control following SRS compared with relying on standard clinical variables.

Complementary information regarding the peritumoral region may be obtained by using a volumetric T2-weighted fluid-attenuated inversion recovery (FLAIR) sequence, which predominantly identifies edema.[1] Several new MRI-based techniques, such as chemical exchange saturation transfer and magnetization transfer,[2,3] show promise for tumor response assessment and prediction; however, they have yet to be widely adopted as standard imaging protocols.

Common treatments for BM include whole brain radiotherapy (WBRT), stereotactic radiosurgery (SRS), surgical resection, and a combination of these modalities. At present, SRS is the dominant treatment choice for patients presenting with a limited number of BM, or as salvage therapy after WBRT. The nature of SRS is to deliver an ablative dose of focal radiation in 1 to 5 fractions to the T1c defined volume. SRS has been reported to be highly effective, with local control rates of 70–80%.[4]

One of the challenges with SRS, particularly for smaller targets, is response determination. Diagnosing local failure and distinguishing it from radiation necrosis is difficult with current conventional imaging. At present, the Response Assessment in Neuro-Oncology (RANO)–BM criteria represent the clinical standard for response determination; however, these guidelines are suboptimal as they are based on a unidimensional assessment and designed to evaluate metastases larger than 10 mm.[5] Given that the treatment of BM has changed such that SRS is routinely delivered for patients with more than 5 lesions (often subcentimeter in diameter), the limitations of RANO-BM (and similar guidelines) to account individual target responses is a significant issue. For example, the notion that the evaluation of a limited number of metastases (up to 5) is a surrogate of overall response in the brain for patients with more than 5 metastases is simply a reflection of limitations in human assessment capability and the current software available in the clinic for diagnostic imaging. In the modern era, there is an urgent need to refine response determination to be based on individual targets as more patients are treated for multiple small subcentimeter lesions. Additionally, it has been observed that the volume of disease is a better predictor of overall survival than the number of lesions.[6,7] Lastly, the biological heterogeneity at the genomic level of individual metastases necessitates that each lesion be accounted for as an independent variable to tailor salvage therapy options.[8] Ultimately, the aim is to develop an accurate method of pretreatment response prediction at the level of the individual metastasis, which holds the potential to improve patient management as identification of BM at risk for local failure may warrant surgery, a higher radiation dose, and/or a change in systemic management.

In recent years, large-scale characterization of tumor phenotype from medical image data, often referred to as radiomics, has shown potential for outcome prediction.[9] Radiomics describes a wide range of computational methods aiming to extract intensity, shape, and texture-based features from a predefined region of interest with the aim of aiding prognosis, diagnosis, and treatment response. Radiologists often derive useful information from tumor appearance on medical images using qualitative descriptors (eg, heterogeneity) or simple quantitative features (eg, maximum diameter, volume).[8] However, qualitative measures are time-consuming, capture few features, and are prone to inter/intraobserver variability.[10] Radiomics addresses several of these limitations by quantitatively characterizing tumor phenotype through computational methods that automatically extract meaningful patterns from imaging data.

Presently, there is limited evidence to suggest that radiomics provides value in predicting local recurrence following SRS for BM based on pretreatment CT and T1c scans.[11–13] There is, however, uncertainty whether radiomics provides additional value beyond clinical, dosimetric, and structural radiographic variables (which we refer to as "clinical features") with respect to outcome prediction. Accordingly, the aim of this study was to determine if radiomic features extracted from standardized volumetric T1c and volumetric FLAIR scans can provide complementary information capable of improving the prediction of local failure of BM after SRS.

Materials and Methods

Patient Data

An institutional research ethics board–approved retrospective review of patients treated for BM with Gamma Knife and Gamma Knife Icon (both Elekta) between December 2016 and November 2017 was performed. Inclusion criteria consisted of histologically confirmed cancer with BM diagnosed on a gadolinium contrast-enhanced MRI scan. Exclusion criteria, for an individual BM: cystic metastases, non-parenchymal metastases, surgical cavities, and metastases having received previous SRS. A total of 408 lesions within 87 SRS patients were included.

Neuro-Oncology

MR Imaging

For each case, pretreatment volumetric T1c and volumetric 3D axial FLAIR scans were obtained on a Philips 1.5T Ingenia system. Image preprocessing consisted of artifact correction, followed by standardization of voxel size, spacing, and intensity. N4 bias field correction was applied to correct for any image artifacts from magnetic field inhomogeneities using the Insight Toolkit (ITK) implementation.[14,15] Hyperintensity artifacts were corrected by suppressing voxel intensities above the 99.9th percentile for each image. Image size and voxel spacing were normalized to the median values in the dataset: (432 × 432 × 180 voxels) and (0.56 × 0.56 × 1.0 mm), respectively, for T1c and (448 × 448 × 160 voxels) and (0.56 × 0.56 × 1.0 mm) for FLAIR, respectively. Image intensity was normalized to have zero mean and unit variance.

Radiomics Feature Extraction

Tumor regions of interest (ROIs) consisted of tumor core and peritumoral volume, as shown in Figure 1. Tumor core was defined by the gross tumor volume (GTV) as delineated by radiation oncologists and reviewed by a neuro-radiologist during treatment planning based on the enhancement on T1c (Figure 1B). Peritumoral volume was calculated by expanding the T1c tumor core mask by half its major axis length using the SimpleITK implementation of the binary dilation morphologic operation,[16,17] then subtracting the original tumor core mask from the expanded region (Figure 1E). These T1c ROIs were mapped onto the corresponding FLAIR image through a transformation matrix, which was obtained via a rigid registration of the T1c and FLAIR images using the Elastix software (Figure 1C, F).[18]

Study Endpoints

Patients were followed as per our institutional protocol with a volumetric MRI scan and clinical follow-up appointment every 2–3 months after SRS. Response status was determined on a lesion-wise basis on each follow-up MRI scan, based on changes in tumor size as indicated by the maximum axial (2D) lesion diameter according to RANO-BM recommendations.[5] However, unlike the RANO-BM criteria, these labels were applied to each individual BM. Lesions that demonstrated an increase in maximum 2D tumor diameter greater than 20%, at any point on the available clinical follow-up MRI scans, were labeled as local failure if they did not subsequently stabilize or regress without further active treatment, in which case they were labeled as adverse radiation effect (ARE). This method is consistent with the literature.[19] All other lesions that demonstrated a RANO-BM stable response, partial response, complete response, or ARE were pooled together as "treatment response." This study did not impose restrictions on the minimum measurable lesion size. Additional information regarding the response assessment procedure used for this study is summarized in Supplementary Tables 1 and 2.

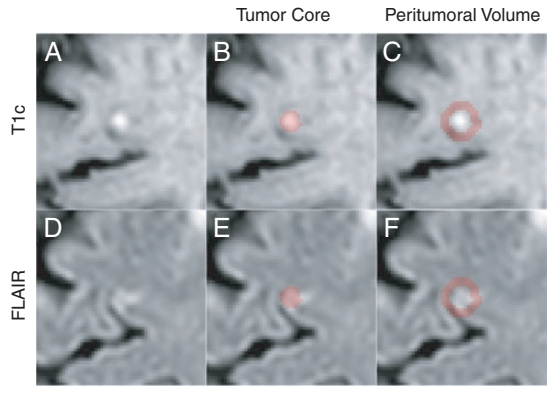

Fig. 1 Axial slice centered on a BM as it appears on T1c (A) and FLAIR (D) scans. Highlighted regions represent the tumor core and peritumoral volume on T1c (B and E, respectively) and FLAIR (C and F, respectively).

Feature Selection and Prediction Modeling

Image features were extracted from the tumor core and peritumoral volume ROIs of the T1c and FLAIR scans (Figure 1). A total of 440 image features were extracted using the Pyradiomics (v2.0.0) package from all ROIs,[20] characterizing first-order statistics, shape, and texture information of the tumor. A summary of each feature class can be found in Supplementary Table 3. Radiomic features were compared in their predictive performance to 6 clinical features: (i) radiation dose specific to the lesion (*Prescription Dose*); (ii) ratio of the prescription dose and maximum dose for the lesion (*Isodose Line Percentage*); (iii) maximum axial tumor diameter on T1c (*Max 2D Diameter*); (iv) number of metastases treated (*Number of Metastases*); (v) location of primary tumor site (*Primary Tumor Site*); and (vi) previous WBRT (*Prev WBRT*). These 6 features were chosen based on availability and accessibility of information of the institutional electronic patient record system and on guidance of a radiation oncologist with the aim of characterizing information currently used for managing the treatment of BM at this institution. Radiomic features were automatically selected by their ability to predict local recurrence, using the random forest feature importance (RFI) metric.[21–24] The dataset was jackknife resampled in a leave-one-patient-out (LOPO) manner—ie, the samples from all patients but one were used to train an RF classifier in order to rank each feature by the importance metric. Top-performing features were selected based on their averaged feature importance rankings across all subsets of the data.

Statistical Analysis

The predictive power of clinical and radiomic information was compared by evaluating the performance of random forest (RF) classifiers trained on selected features. Model performance was evaluated in a similar LOPO manner with the number of cross-folds corresponding to the number of patients. The validation set within each cross-fold consisted of samples from a single patient, while the

rest of the samples were used for training. Since each patient had 3 metastases on average treated with SRS, there were relatively fewer samples within the validation set of each cross-fold, which would result in high variance when evaluating the classifier. In order to address this, a single validation set was constructed by concatenating the predictions made on the validation samples within each cross-fold. Predictions generated on the samples of the concatenated validation set were bootstrap resampled 6000 times with replacement, and stratified by endpoints. The area under the receiver operating characteristic curve (AUC) was calculated for each of the 6000 subsamples, with the mean AUC and 95% confidence interval of this mean characterizing the model performance and uncertainty, respectively. Statistical comparisons between 2 classifiers trained on different features were made using a paired *t*-test of AUC values from the resampled dataset.[25,26]

Results

A total of 440 radiomic features and 6 clinical features were compared in their ability to predict local lesion control on a dataset of 408 lesions from 87 patients. A summary of patient and lesion characteristics is provided in Table 1. The median follow-up imaging duration was 6.1 months (range, 0.7–17.6). For those patients demonstrating progression/ARE, the additional median follow-up time interval was 1.9 months (range, 0.0–12.0) with a median of one additional MR scan performed. A total of 3 lesions were suspected of ARE. An 8% crude rate of tumor progression was observed. The Kaplan–Meier plot for progression-free survival probability with the horizontal/time axis representing the duration from the treatment date to a progressive disease diagnosis (event), or last available follow-up (non-event), is shown in Figure 2A. The distribution of tumor progression labels based on the RANO-BM response criteria at the last available follow-up, excluding lesions suspected of ARE, is demonstrated in Figure 2B.

Radiomic features were ranked on their ability to predict local control using the RFI metric. A single top-ranked radiomic feature was combined with the set of clinical features and compared with the set of clinical features alone (control). Out of the top 10 ranked radiomic features, the addition of any single one resulted in a significant ($P < 0.001$) increase in mean resampled AUC (Table 2). The highest increase in performance from the addition of a single radiomic feature was found with *T1c_peritumoralVolume_firstorder_90Percentile*, which resulted in mean resampled AUC of 0.718 (95% CI of mean: 0.717–0.720) compared with clinical features alone, with an AUC of 0.669 (0.668–0.671) (Table 2).

To gauge the upper limits on the potential increase in predictive power from incorporating radiomic information, all features (440 radiomic and 7 clinical) were combined and ranked based on feature importance as before. An additional step was taken to filter highly correlated features ($R^2 > 0.7$) to further reduce redundancy (Figure 3A). The top-performing features were recursively added for use in training the classifier in order of their importance

Table 1 Patient and lesion characteristics

Variable	Data
Number of Patients	87
Median age (years), range	63, 34–88
Median follow-up (months), range	6.1, 0.7–17.6
Female patients (% patients)	52 (60%)
Male patients (% patients)	35 (40%)
Median number of metastases per patient, range	3, 1–33
No prior WBRT (% patients)	72 (83%)
Prior WBRT (% patients)	14 (16%)
N/A prior WBRT (% patients)	1 (1%)
Primary tumor site: No. Failures, No. Lesions, Crude local failure rate (% of all lesions)	
Breast	0, 53, 0.0%
Lung	14, 202, 6.9%
Other	4, 42, 9.5%
Melanoma	14, 111, 12.6%
Total number of metastases	408
Median tumor volume (cc), range	0.12, 0.01–8.9
Lesion crude local failure (No. lesions), % of all lesions	32, (7.8%)
Maximum 2D diameter – *d*: No. Lesions (Local failure rate)	
$0 < d \leq 5$ mm	115 (5.2%)
$5 < d \leq 10$ mm	184 (4.9%)
$10 < d \leq 20$ mm	93 (14.0%)
$d \geq 20$ mm	16 (25%)
Median prescribed dose (Gy), range	20, 14–25
Median isodose line percentage (% of lesions), range	56, 41–95

Abbreviations: No. – Number of. N/A—Not Available. cc—cubic centimeters.

rank (Table 3). A combination of the top 12 highest ranked features resulted in the highest mean AUC of 0.793 (0.792–0.795) (Figure 3B), which was 19% higher than the performance of clinical features alone.

In order to gain more insight into each individual feature, univariate analysis was performed on each of the top-ranked radiomic and clinical features by comparing local control and local failure groups. Each of the top-ranked clinical features shows a significant difference, while only 2 radiomics features (*T1c_tumorCore_shape_Sphericity*, *T1c_tumorCore_glrlm_ShortRunEmphasis*) were significant on univariate analysis. Further analysis showed that *T1c_tumorCore_glrlm_ShortRunEmphasis* is highly correlated with the logarithm of *T1c_tumorCore_shape_Volume* (), while *T1c_tumorCore_shape_Sphericity* demonstrates a much weaker linear () and logarithmic () relationship with *T1c_tumorCore_shape_Volume*. A similar comparison between top-ranked clinical features and volumetric/unidimensional measurements of tumor core size did not reveal a strong linear or log-linear relationship ().

Discussion

We report that incorporating radiomic features, alongside clinical features, improves local response prediction for BM treated with SRS. Each of the top 10 best performing radiomic features demonstrated significantly improved predictive performance when combined with clinical features. An optimized combination of radiomic and clinical features further improved the mean AUC. These results suggest that quantitative characterization of tumor structure through the extraction of radiomic features from T1c

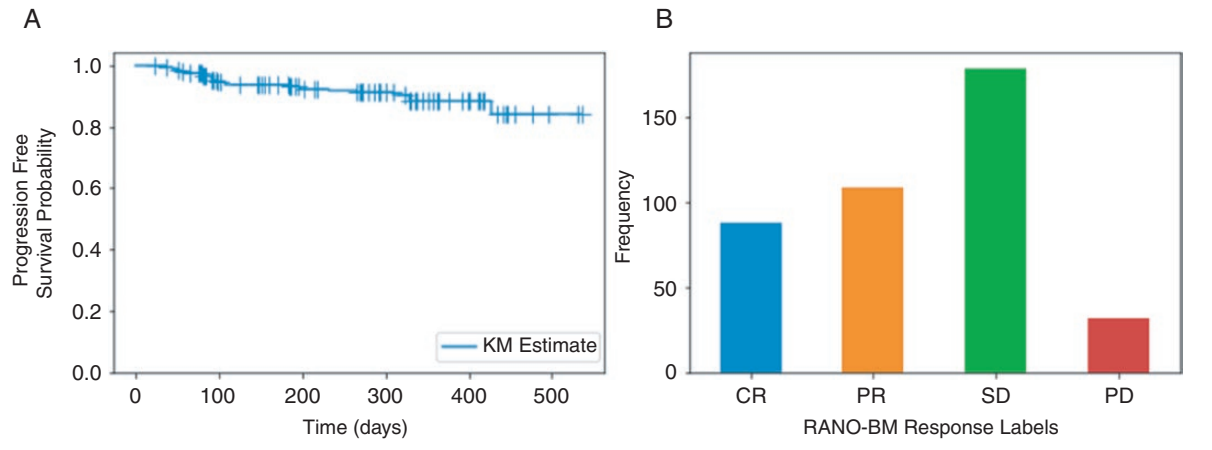

Fig. 2 (A) Kaplan–Meier curve for progression-free survival. (B) Bar plot of progression labels based on the RANO-BM response criteria for target lesions at the last available follow-up scan (excluding patients suspected of ARE). += Censored. CR—Complete Response. PR—Partial Response. SD—Stable Disease. PD—Progressive Disease.

Table 2 Comparison of AUC values between clinical features alone and clinical features with a single top-performing radiomic feature

Feature Set	AUC† (95% CI of mean)	*P*-value†† (V. Baseline)
Clinical Features (Baseline)	0.669 (0.668, 0.671)	NA
Clinical Features + T1c_tumorCore_shape_Sphericity	0.690 (0.688, 0.691)	<0.001 *
Clinical Features + T1c_peritumoralVolume_firstorder_Maximum	0.693 (0.692, 0.695)	<0.001 *
Clinical Features + T1c_peritumoralVolume_firstorder_90Percentile	0.718 (0.717, 0.720)	<0.001 *
Clinical Features + T1c_tumorCore_shape_Flatness	0.694 (0.693, 0.696)	<0.001 *
Clinical Features + T1c_peritumoralVolume_firstorder_Range	0.680 (0.678, 0.682)	<0.001 *
Clinical Features + T1c_tumorCore_glrlm_ShortRunEmphasis	0.692 (0.690, 0.694)	<0.001 *
Clinical Features + T1c_peritumoralVolume_firstorder_RootMeanSquared	0.681 (0.679, 0.683)	<0.001 *
Clinical Features + T1c_tumorCore_shape_Elongation	0.685 (0.684, 0.687)	<0.001 *
Clinical Features + FLAIR_tumorCore_firstorder_Range	0.704 (0.702, 0.706)	<0.001 *
Clinical Features + T1c_tumorCore_firstorder_Maximum	0.682 (0.680, 0.683)	<0.001 *

Note: AUC scores were generated from the predictions of the RF classifier on the concatenated validation set. Radiomics features are listed in order of decreasing feature importance rank.
Abbreviations: * Statistically significant. † mean resampled validation AUC (6000 times). †† One-sided *P*-value.

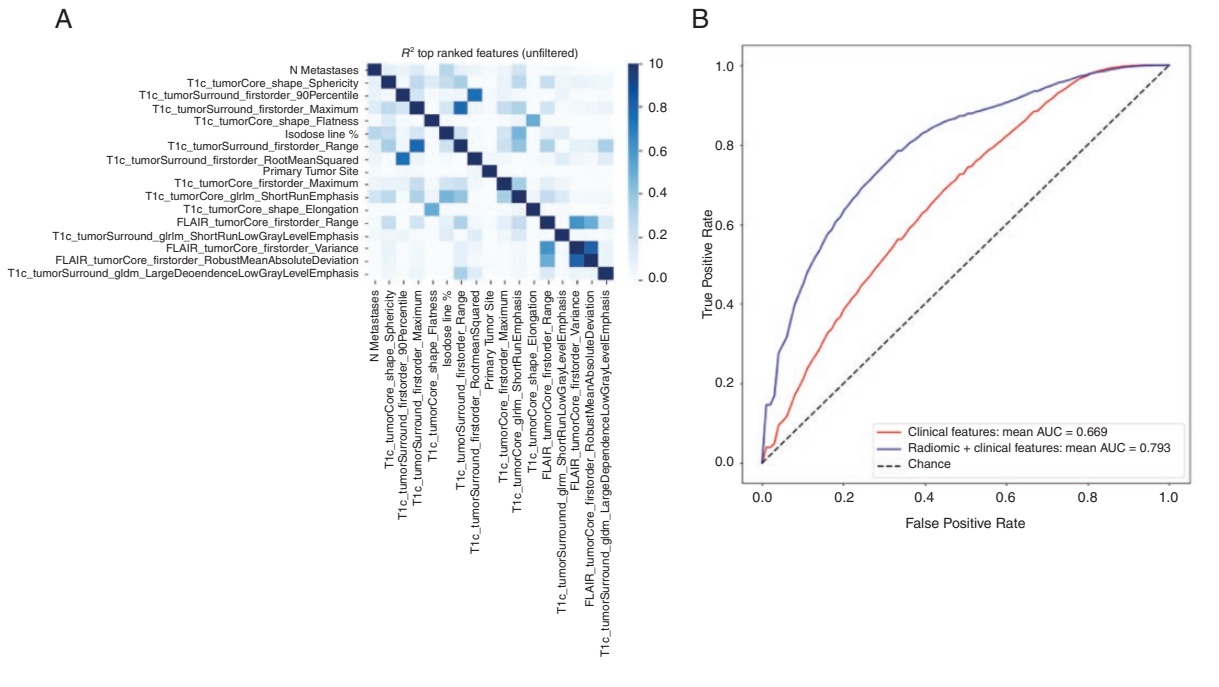

Fig. 3 (A) Heatmap of pairwise Pearson correlation coefficients for top-ranked clinical and radiomic features. (B) Average ROC curve obtained from resampling the predictions made on the concatenated validation set. The red (blue) curve corresponds to the RF model trained on clinical (imaging and clinical) features.

and FLAIR scans adds complementary information to what is typically available in the clinical workflow.

Determining a biological mechanism driving the predictive value of biomarkers is an active challenge in the field of radiomics. Due to the feature selection method used in this study, which measured the average drop in performance if the feature is removed, the features determined to be important are not necessarily predictive of local failure on their own. Thus, it is difficult to assign any definitive meaning to the radiomic signature on a feature-wise basis. Although other methods, such as logistic regression, provide a more intuitive sense of each feature's individual value, they lack the ability to utilize interactions between features necessary to maximize the predictive power of the radiomic signature as a whole.

Two radiomic features that emerged as both important (based on RFI) and significant on univariate analysis were *T1c_tumorCore_glrlm_ShortRunEmphasis* and *T1c_tumorCore_shape_Sphericity*. Further analysis showed that *T1c_tumorCore_glrlm_ShortRunEmphasis* holds a log-linear relationship with tumor core volume on T1c () and is likely a surrogate for tumor size, which is already a known predictor of local failure.[7] *T1c_tumorCore_shape_Sphericity* appears to indicate a higher probability of local failure for low values of this feature and only holds a weak relationship to tumor volume, suggesting that it is an independent feature. Justification for why tumor sphericity was found important may perhaps be explained by an invasive cancer phenotype giving rise to spatial variations at the border between healthy tissue and tumor, resulting in lower values of sphericity. This feature was also shown to be predictive of tumor recurrence and survival in patients with oral cavity squamous cell carcinoma and pathological response in patients with non-small-cell lung carcinoma, which further supports its prognostic value.[26,27] Visualization of the tumor sphericity feature may be seen in Supplementary Figure 1.

In this study, we used the RFI to determine top-performing radiomic and clinical features. Features that were found to be important using this metric were not necessarily significant on univariate analysis. This is because RFI is a multivariate technique which can gauge the effectiveness of each feature in the presence of other features (ie, accounts for feature interactions), unlike the Mann–Whitney *U*-test, which was used for univariate analysis. While we believe that RFI is a better metric to use in feature selection, it does not inform on the importance of any individual feature by itself. For this reason, we used univariate analysis to determine if any single feature has a direct independent relationship with local control. This secondary analysis informs why certain features improve the classifier's performance and what mechanisms may be driving local failure.

Out of the top-ranked clinical features, *Number of Metastases Treated, Isodose Line Percentage*, and *Primary Tumor Site* were significant on univariate analysis. A lower value for the *Number of Metastases* corresponded to a higher risk of local failure. This was based on the number of brain metastases per patient. It is likely explained by the fact that patients with multiple metastases had smaller lesions and the existence of an inherent patient selection bias

Neuro-Oncology

Table 3 AUC scores of RF classifier using top *n* ranked radiomic and clinical features

Feature Name	Importance Rank	AUC† (95% CI of mean)	Univariate Analysis *P* -Value††
Number of Metastases	1	0.545 (0.543, 0.547)	<0.01 *
T1c_tumorCore_shape_Sphericity	2	0.665 (0.663, 0.667)	<0.001 *
T1c_peritumoralVolume_firstorder_90Percentile	3	0.762 (0.760, 0.763)	0.21
T1c_peritumoralVolume_firstorder_Maximum	4	0.770 (0.769, 0.772)	0.26
T1c_tumorCore_shape_Flatness	5	0.781 (0.780, 0.783)	0.11
Isodose line Percentage	6	0.780 (0.779, 0.782)	<0.005 *
Primary Tumor Site	7	0.786 (0.784, 0.787)	<0.01 *
T1c_tumorCore_firstorder_Maximum	8	0.776 (0.775, 0.778)	0.14
T1c_tumorCore_glrlm_ShortRunEmphasis	9	0.777 (0.776, 0.779)	<0.001 *
T1c_tumorCore_shape_Elongation	10	0.778 (0.777, 0.780)	0.07
FLAIR_tumorCore_firstorder_Range	11	0.775 (0.774, 0.777)	0.20
T1c_peritumoralVolume_glrlm_ ShortRunLowGrayLevelEmphasis	12	0.793 (0.792, 0.795)	0.43
FLAIR_tumorCore_firstorder_Variance	13	0.779 (0.778, 0.781)	0.26
T1c_peritumoralVolume_gldm_ LargeDependenceLowGrayLevelEmphasis	14	0.778 (0.777, 0.780)	0.09

Note: Features ranked 1 through k were used for training of RF Classifier. Features are listed in order of decreasing feature importance rank.
Abbreviations: † mean resampled AUC (6000 times). †† Univariate Mann–Whitney *U*-test (one-sided). * Significant.

associated with technical delivery and safety. For example, one patient had 33 brain metastases, of which 76% were <5 mm. Regarding *Isodose Line Percentage*, we observed that local control was greater when lesions were treated with a higher prescription isodose line. This is likely a reflection of an inherent margin of therapeutic dose around the metastasis, as the dose fall-off is less steep when prescribing to a higher isodose line. The typical treatment with SRS includes no margin beyond the contrast-enhancing lesion, despite data supporting a 1 mm safety margin beyond the GTV resulting in higher local control without an increase in complication rates.[28,29] Therefore, the practice of prescribing to a higher percentage isodose line allows for essentially more generous coverage of potential microscopic disease. This result is also consistent with a report by Sheehan et al of a cohort of lung cancer patients with BM treated with SRS.[30] Interestingly, this predictor and tumor sphericity may be related and assist in explaining this radiomic feature. Lastly, *Primary Tumor Site* was also observed to be predictive. This likely reflects the greater risk of relapse observed in this cohort within the melanoma and lung cancer primary tumor types compared with the number of failures in the breast cancer patients.[31]

Out of the top 10 radiomic features, our feature selection method only identified one important FLAIR-based feature (*FLAIR_tumorCore_firstorder_Range*) compared with 9 T1c-based features. A similar lack of important FLAIR-based features is seen when feature importance is gauged with the presence of both clinical and radiomic features. We used a simple automated approach to define the FLAIR ROIs, which consisted of expanding the region around the enhancing tumor by a uniform amount. This approach may not capture all the FLAIR hyperintensity, and may also include healthy brain tissue, leading to noise being included in the FLAIR-related radiomic features. Further work is needed to determine whether a more accurate segmentation of the FLAIR hyperintensity would result in an increase in importance of the FLAIR radiomic features.

The sample size is a limiting factor of this study, and a larger dataset would be required for independent feature selection, increased accuracy, greater generalizability, and a stratified analysis of individual features (eg, by tumor histology, age, sex, etc). Another limitation of our study, and of many similar studies, was the determination of lesion endpoints. Due to low resection rates of BM in this cohort upon progression, local failure was determined radiographically based on a unidimensional measure of tumor size according to RANO-BM criteria. It is acknowledged that an increase in tumor dimensions can represent ARE as opposed to true tumor progression, and vice versa, and it is why we are in need of non-invasive yet reliable means to image tumor viability post-SRS.[32] Although lesions that showed progression and then stabilized without further treatment were subsequently labeled as ARE, it is acknowledged that histopathologic analysis is the gold standard to determine ARE versus tumor progression. It is not feasible to resect or biopsy most questionable lesions posttreatment unless clinically indicated (such as those causing symptomatic edema).[33,34] Future work should focus on collecting a larger multi-institutional dataset, with expert ROI annotations on both T1c and FLAIR scans, which includes histopathologic analyses of resected brain metastases.

Although our image data were limited to a single homogeneous scan sequence from a single institution, this resulted in reduced variability of the radiomic features and consequently allowed for more powerful inferences. Another

strength of this study was in the lesion segmentation procedure. In order to ensure accurate ROIs for extraction of robust radiomics features, the T1c tumor core ROIs were obtained from GTV treatment planning contours delineated by a radiation oncologist. Lastly, the inclusion of volumetric FLAIR scans in our pipeline provided an additional source of radiomic data and consequently, a new set of image features that are not present in similar radiomics studies.[11–13]

Conclusion

We have shown that the addition of radiomic features provides complementary information to standard routinely available clinical variables for the prediction of local failure in BM after SRS. A predictive model based on information provided by radiomic and clinical features shows promise for pretreatment outcome prediction, and may also prove useful as *a priori* information when differentiating ARE from tumor progression for lesions that demonstrate an increase in size on T1c during follow-up. The measure of tumor sphericity on T1c warrants further investigation on new patient cohorts, as it may contain valuable information for understanding why some BM fail after SRS.

Supplementary Material

Supplementary data are available at *Neuro-Oncology* online.

Keywords

brain metastases | radiomics | response prediction | neuro-oncology | local control

Funding

This research was supported by the Sunnybrook Research Institute, through funding from the Federal Economic Development Agency for Southern Ontario (FedDev Ontario). Graduate student funding comes in part from the Natural Sciences and Engineering Research Council (NSERC).
An abstract summarizing the results of this article has been accepted as a poster presentation at the 2019 ASTRO conference.

Conflict of interest statement.
Dr Ruschin:
- Co-inventor of and owns associated intellectual property specific to the image-guidance system on the Gamma Knife Icon.

Dr Sahgal:
- Advisor/consultant with Abbvie, Merck, Roche, Varian (Medical Advisory Group), Elekta (Gamma Knife Icon), BrainLAB, and VieCure (Medical Advisory Board)
- Board Member: International Stereotactic Radiosurgery Society (ISRS)
- Past educational seminars with Elekta AB, Accuray Inc, Varian (CNS Teaching Faculty), BrainLAB, Medtronic Kyphon
- Research grant with Elekta AB
- Travel accommodations/expenses by Elekta, Varian, BrainLAB
- Dr Sahgal also belongs to the Elekta MR Linac Research Consortium, Elekta Spine, Oligometastases and Linac Based SRS Consortia

The rest of the authors have no conflict of interest to disclose.

Authorship statement. *Experimental design*: Anne L. Martel, Mark Ruschin, Chris Heyn, Irene Karam, Young K. Lee, Jay Detsky. *Implementation*: Andrei Mouraviev, Anne Martel. *Analysis*: Anne Martel, Andrei Mouraviev. *Interpretation*: Jay Detsky, Arjun Saghal, Anne Martel, Greg J. Stanisz. *Writing and revision of manuscript*: All authors.

References

1. Drevelegas A, Papanikolaou N. Imaging modalities in brain tumors. In: *Imaging of Brain Tumors with Histological Correlations*. Berlin, Heidelberg, Germany: Springer; 2011:13–33.
2. Desmond KL, Mehrabian H, Chavez S, et al. Chemical exchange saturation transfer for predicting response to stereotactic radiosurgery in human brain metastasis. *Magn Reson Med*. 2017;78(3):1110–1120.
3. Mehrabian H, Myrehaug S, Soliman H, Sahgal A, Stanisz GJ. Quantitative magnetization transfer in monitoring glioblastoma (GBM) response to therapy. *Sci Rep*. 2018;8(1):2475.
4. Chao ST, De Salles A, Hayashi M, et al. Stereotactic radiosurgery in the management of limited (1–4) brain metastases: systematic review and International Stereotactic Radiosurgery Society practice guideline. *Clin Neurosurg*. 2018;83(3):345–353.
5. Lin NU, Lee EQ, Aoyama H, et al. Response assessment criteria for brain metastases: proposal from the RANO group. *Lancet Oncol*. 2015;16(6):e270–e278.
6. Sahgal A, Ruschin M, Ma L, Verbakel W, Larson D, Brown PD. Stereotactic radiosurgery alone for multiple brain metastases? A review of clinical and technical issues. *Neuro Oncol*. 2017;19(suppl_2):ii2–ii15.
7. Follwell MJ, Khu KJ, Cheng L, et al. Volume specific response criteria for brain metastases following salvage stereotactic radiosurgery and associated predictors of response. *Acta Oncol (Madr)*. 2012;51(5):629–635.
8. Brastianos PK, Carter SL, Santagata S, et al. Genomic characterization of brain metastases reveals branched evolution and potential therapeutic targets. *Cancer Discov*. 2015;5(11):1164–1177.
9. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology*. 2016;278(2):563–577.
10. Bauknecht H-C, Romano VC, Rogalla P, et al. Intra- and interobserver variability of linear and volumetric measurements of brain metastases using contrast-enhanced magnetic resonance imaging. *Invest Radiol*. 2010;45(1):49–56.
11. Cha Y, Kim MS, Cho CK, et al. Prediction of local response after stereotactic radiosurgery for brain metastases using convolutional neural networks based radiomics. *Int J Radiat Oncol*. 2017;99(2):E645.

12. Yeung TPC, Rodrigues G, Lagerwaard F, et al. Prediction of stereotactic radiosurgery brain metastasis lesion control using radiomic features. *Int J Radiat Oncol.* 2015;93(3):S7.
13. Wang H, Barbee D, Xue J, Das IJ, Kondziolka D. Predicting local recurrence of stereotactic radiosurgery brain metastases using MRI radiomics features. *Int J Radiat Oncol.* 2018;102(3):e563–e564.
14. Tustison NJ, Avants BB, Cook PA, et al. N4ITK: improved N3 bias correction. *IEEE Trans Med Imaging.* 2010;29(6):1310–1320.
15. Ibáñez L, Ibá I, Ibáñez I, Schroeder W, Ng L, Cates J. The ITK Software Guide.; 2003. http://www.itk.org. Accessed September 13, 2019.
16. Lowekamp BC, Chen DT, Ibáñez L, Blezek D. The design of SimpleITK. *Front Neuroinform.* 2013;7:45.
17. Nikopoulos N, Pitas I. An efficient algorithm for 3D binary morphological transformations with 3D structuring elements of arbitrary size and shape. *Proceedings of 1997 IEEE Workshop on Nonlinear Signal and Image Processing (NSIP'97*). 1997.
18. Klein S, Staring M, Murphy K, Viergever MA, Pluim JP. elastix: a toolbox for intensity-based medical image registration. *IEEE Trans Med Imaging.* 2010;29(1):196–205.
19. Sneed PK, Mendez J, Vemer-van den Hoek JG, et al. Adverse radiation effect after stereotactic radiosurgery for brain metastases: incidence, time course, and risk factors. *J Neurosurg.* 2015;123(2):373–386.
20. van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res.* 2017;77(21):e104–e107.
21. Louppe G. Understanding Random Forests: From Theory to Practice. July 2014. http://arxiv.org/abs/1407.7502. Accessed September 13, 2019.
22. Pedregosa F, Michel V, Grisel O, et al. Scikit-learn: machine learning in python. *J Machine Learn Res.* 2011;12(OCT):2825–2830. http://scikit-learn.sourceforge.net. Accessed September 13, 2019.
23. Breiman L. Random forests. *Mach Learn.* 2001;45(1):5–32.
24. Ho TK. Random decision forests. In: Proceedings of the International Conference on Document Analysis and Recognition, ICDAR. Vol 1. IEEE Computer Society; 1995:278–282.
25. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics.* 2011;12:77.
26. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology.* 1983;148(3):839–843.
27. Tarsitano A, Ricotta F, Cercenelli L, et al. Pretreatment tumor volume and tumor sphericity as prognostic factors in patients with oral cavity squamous cell carcinoma. *J Cranio-Maxillofacial Surg.* 2019;47(3):510–515
28. Coroller TP, Agrawal V, Huynh E, et al. Radiomic-based pathological response prediction from primary tumors and lymph nodes in NSCLC. *J Thorac Oncol.* 2017;12(3):467–476.
29. Noël G, Simon JM, Valery CA, et al. Radiosurgery for brain metastasis: impact of CTV on local control. *Radiother Oncol.* 2003;68(1):15–21.
30. Sheehan JP, Sun MH, Kondziolka D, Flickinger J, Lunsford LD. Radiosurgery for non-small cell lung carcinoma metastatic to the brain: long-term outcomes and prognostic factors influencing patient survival time and local tumor control. *J Neurosurg.* 2002;97(6):1276–1281.
31. Chang EL, Selek U, Hassenbusch SJ 3rd, et al. Outcome variation among "radioresistant" brain metastases treated with stereotactic radiosurgery. *Neurosurgery.* 2005;56(5):936–45; discussion 936.
32. Patel TR, McHugh BJ, Bi WL, Minja FJ, Knisely JP, Chiang VL. A comprehensive review of MR imaging changes following radiosurgery to 500 brain metastases. *AJNR Am J Neuroradiol.* 2011;32(10):1885–1892.
33. Detsky JS, Keith J, Conklin J, et al. Differentiating radiation necrosis from tumor progression in brain metastases treated with stereotactic radiotherapy: utility of intravoxel incoherent motion perfusion MRI and correlation with histopathology. *J Neurooncol.* 2017;134(2):433–441.
34. Szeifert GT, Atteberry DS, Kondziolka D, Levivier M, Lunsford LD. Cerebral metastases pathology after radiosurgery: a multicenter study. *Cancer.* 2006;106(12):2672–2681.