Invited Original Study

OPEN

Radiomics-Based Prediction of Long-Term Treatment Response of Vestibular Schwannomas Following Stereotactic Radiosurgery

*†Patrick P. J. H. Langenhuizen, †Svetlana Zinger, ‡Sieger Leenstra, §||Henricus P. M. Kunst, §Jef J. S. Mulder, *Patrick E. J. Hanssens, †Peter H. N. de With, and *Jeroen B. Verheul

*Gamma Knife Center, Department of Neurosurgery, ETZ Hospital, Tilburg; †Eindhoven University of Technology, Eindhoven; ‡Department of Neurosurgery, Erasmus Medical Center, Rotterdam; §Department of Otolaryngology, Radboud University Medical Center, Radboud Institute of Health Sciences, Nijmegen; and ||Departments of Otolaryngology, Head and Neck Surgery, and Neurosurgery, Maastricht University Medical Centre, Maastricht, the Netherlands

Objective: Stereotactic radiosurgery (SRS) is one of the treatment modalities for vestibular schwannomas (VSs). However, tumor progression can still occur after treatment. Currently, it remains unknown how to predict long-term SRS treatment outcome. This study investigates possible magnetic resonance imaging (MRI)-based predictors of long-term tumor control following SRS.
Study Design: Retrospective cohort study.
Setting: Tertiary referral center.
Patients: Analysis was performed on a database containing 735 patients with unilateral VS, treated with SRS between June 2002 and December 2014. Using strict volumetric criteria for long-term tumor control and tumor progression, a total of 85 patients were included for tumor texture analysis.
Intervention(s): All patients underwent SRS and had at least 2 years of follow-up.
Main Outcome Measure(s): Quantitative tumor texture features were extracted from conventional MRI scans. These features were supplied to a machine learning stage to train prediction models. Prediction accuracy, sensitivity, specificity, and area under the receiver operating curve (AUC) are evaluated.
Results: Gray-level co-occurrence matrices, which capture statistics from specific MRI tumor texture features, obtained the best prediction scores: 0.77 accuracy, 0.71 sensitivity, 0.83 specificity, and 0.93 AUC. These prediction scores further improved to 0.83, 0.83, 0.82, and 0.99, respectively, for tumors larger than 5 cm^3.
Conclusions: Results of this study show the feasibility of predicting the long-term SRS treatment response of VS tumors on an individual basis, using MRI-based tumor texture features. These results can be exploited for further research into creating a clinical decision support system, facilitating physicians, and patients to select a personalized optimal treatment strategy. **Key Words:** Machine learning—Magnetic resonance imaging—Radiomics—Stereotactic radiosurgery—Treatment prediction—Tumor texture—Vestibular schwannoma.

Otol Neurotol **41:**e1321–e1327, 2020.

The main treatment goal for vestibular schwannomas (VSs) has shifted in the last decades from complete removal of the tumor to functional preservation of adjoining cranial nerves in cases where the tumor is not life-threatening (1). Consequently, stereotactic radiosurgery (SRS) is considered to be an attractive alternative of microsurgery, especially for small- to medium-sized VSs (2). The motivation for this is found in lower overall posttreatment morbidity and overall increased preservation rates of cranial nerve functions when compared with microsurgery (3). Moreover, microsurgery invokes a substantially larger overall cost on average, compared with SRS (4,5). For large VSs, utilizing SRS as treatment strategy is significantly more controversial. If clinical symptoms due to their mass effect are present, microsurgery is the obvious choice. However, if these symptoms are absent, there is no clear volume threshold for considering microsurgery over SRS, and each medical center has its own strategy concerning larger VSs. Most medical centers consider microsurgical resection as the optimal treatment strategy (6–8). However, there are studies showing the safety and efficacy of SRS for treating large

Address correspondence and reprint requests to Patrick P. J. H. Langenhuizen, M.Sc., Signal Processing Systems, Department of Electrical Engineering, Eindhoven University of Technology, Groene loper 19, 5612 AP Eindhoven, The Netherlands; E-mail: p.p.j.h.langenhuizen@tue.nl; p.langenhuizen@etz.nl

Conflicts of Interest and Source of Funding: P. P. J. H. Langenhuizen is currently receiving a grant (#80-84200-98-15222) from the Netherlands organization for Health Research and Development.
For the remaining authors none were declared.
The authors disclose no conflicts of interest.
Supplemental digital content is available in the text.

DOI: 10.1097/MAO.0000000000002886

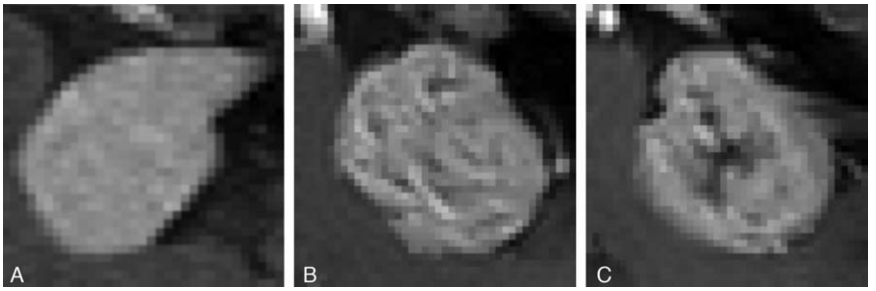

FIG. 1. Examples of different vestibular schwannoma tumor textures on contrast-enhanced MRI scans. The depicted tumors have comparable volumes. Part *A*: near-homogeneously enhanced lesion. Part *B*: small irregularities in texture. Part *C*: Heterogeneously enhanced lesion with an apparent hypo-intense area. MRI indicates magnetic resonance imaging.

and giant VS tumors (7–14). Furthermore, with increasing tumor size, microsurgery becomes more complicated and the risk of surgery inducing morbidity increases (15–17). This makes the potential benefit of SRS even more relevant for larger tumors.

However, SRS on VSs has a significant drawback: long-term tumor control is not obtained in all patients. If tumor progression is not stopped, salvage microsurgical treatment may be needed. Such an intervention negates the previously mentioned reasons for electing SRS over microsurgery in the first place. Moreover, microsurgical excision of tumor tissue following SRS is considered more difficult than if the tumor was not irradiated (18). Thus, it can be concluded that SRS is an attractive treatment modality for VSs, but only if the tumor responds well to this treatment.

Currently, it is not possible to a priori predict the SRS treatment response of a VS on an individual basis. To enable such prediction, tumor-specific information should be assessed. Readily accessible clinical data that can provide such information are available in the form of magnetic resonance imaging (MRI) scans. It is well-known that the MRI findings in VSs are highly variable in the gray-level inhomogeneity of the tissue itself. VS tumors can appear micro- or macro-cystic (19), hemorrhagic (20), and with variable contrast enhancement patterns. Some examples can be found in Figure 1. These MRI differences reflect variations in histology, such as cell proliferation and microvessel density (21,22). As such, these MRI images may provide sufficient information to enable the individual prediction of the SRS treatment response.

In this paper we therefore investigate quantitative, tumor-specific parameters obtained from conventional MRI scans. These so-called radiomic features may provide information on differences in tumor biology, enabling the creation of a patient-specific tumor model that can be employed for predicting the long-term SRS treatment response. The aim of this study is to explore the prediction of long-term tumor control, employing radiomic features obtained from MRI scans.

MATERIALS AND METHODS

The medical ethics review committee waived a formal approval procedure for this retrospective study.

Patient Cohort

All patients with unilateral VS treated with SRS in our center between 2002 and 2014 were identified. This cohort consisted of all VS patients remaining after excluding patients with neurofibromatosis type 2 (NF2), those previously treated for their VS, or with less than 2 years of post-SRS follow-up. Furthermore, we excluded patients with small VS tumors, as these tumors show little to no variation in texture. The associated volumes of such small tumors are ill-defined. Analysis of the voxel intensity variations in our data showed that tumor texture becomes discernible in tumors around $1\,cm^3$ (unpublished data). Since an exact volumetric cut-off is arbitrary, we opted for an approach based on the controversy to treat larger tumors with SRS. Thus, we investigated all tumors with a minimum volume of $1.42\,cm^3$, because this corresponds to the volumes reported in literature for Koos Grade IV tumors (23).

A widely applied approach for obtaining models that are able to predict treatment outcomes, is supervised machine learning (sML) (24). This technique employs training data with predetermined classification labels to discover and identify specific patterns, possibly undiscernible to the human eye, that distinguishes cases in the training data with correct labels. As such, it is crucial that the input data has high-quality labels (i.e., a high certainty of the assigned classification label) such that the trained model is robust. Therefore, a well-defined classification labeling is needed in this research. These labels were defined as long-term tumor control, and true tumor progression. Consequently, strict definitions for both controlled and progressive tumors are required and are discussed below. Patients that could not be labeled were excluded from further analysis.

Treatment and Follow-Up

Stereotactic radiosurgery was performed using the Leksell Gamma Knife model 4C or Perfexion (since November 2008; both Elekta AB, Stockholm, Sweden). A dose of 13 Gy was prescribed to the isodose line, covering 90 to 99% of the tumor volume. For treatment planning, T1-weighted with (T1CE) and without (T1) Gadolinium administration, and T2-weighted (T2) MRI scans were obtained for each patient. Treatment was carried out in a single fraction with frame-based fixation.

After treatment, each patient was subjected to a follow-up schedule with a standard interval of 1 year. A T1CE MRI with a slice thickness of 1 mm was obtained at each follow-up visit. In the case of suspected radiological progression or new or worsening symptoms, the standard interval was reduced. If the tumor displayed radiological regression or stability for several years, the standard interval was extended. All follow-up MRIs were employed for determining a volumetric treatment response.

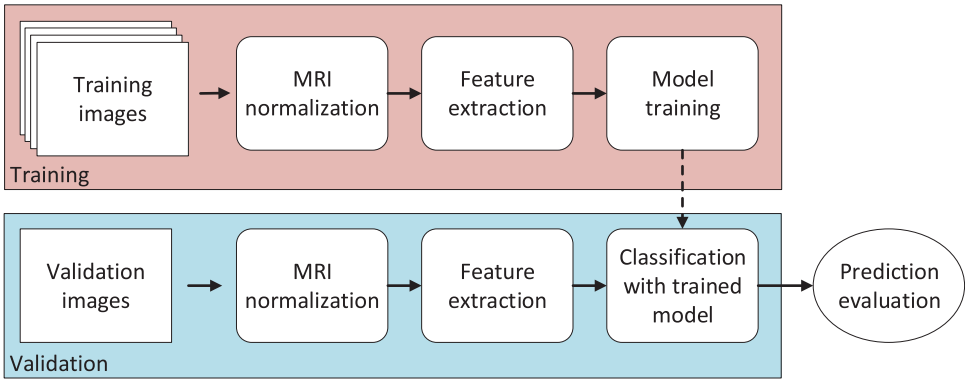

FIG. 2. Processing block diagram of the development of a prediction model.

Tumor volumes were determined by segmenting the tumor using the treatment planning software (GammaPlan versions 10 and 11, Elekta AB, Stockholm, Sweden).

Definitions of Treatment Outcome

True tumor progression was detected using linear measurements in a clinical setting. An increase in tumor size was accepted and considered as radiation-induced swelling during the first 2 years after treatment (25,26), unless the enlargement was deemed too excessive for the considered patient by the radiosurgical team. These failures may have been the result from swelling and not due to true tumor progression. As radiation-induced swelling is a radiobiological distinct response from true tumor progression, we excluded patients that had salvage treatment within the first 2 years following SRS, to avoid inaccuracy through misclassification. Volumetric progression after this period was considered true tumor progression. This was always confirmed by the radiosurgical team. In addition, we looked for discrete volume increases after the initial 2 years, which were undetected by linear measurements performed in the clinical setting. However, these increases were detectable with the volumetric analyses performed in this research. These potentially missed failures are defined as two consecutive significant increases in tumor volume among three consecutive follow-up MRIs, where a minimum of 10% increase in volume is deemed significant (27).

A definition for long-term tumor control highly depends on the time-period in which treatment failures could still occur. There is no certainty that a specific VS will not progress anymore after a predefined time. Hence, long-term tumor control cannot be defined without concessions. Therefore, in this study, we defined it as absence of progression beyond 129 months following treatment. This cutoff is based on the latest-occurring failure in our large database, which was identified at 129 months after treatment.

Post-Processing and Model Training

To enable individual treatment outcome prediction, an sML algorithm is constructed that analyzes all MRI voxels within the tumor contour. Figure 2 depicts a general block diagram of such an algorithm. The different stages of this algorithm are described below.

First, for each patient, the T1-, T1CE-, and T2-MRI scans (Intera and Ingenia, both Philips Healthcare, Best, the Netherlands) were extracted from the treatment planning system, including tumor contours drawn by the neurosurgeon during treatment planning. The matrix sizes of these MRIs were 256×256 for T1 and T1CE, and 512×512 for T2.

Second, the MRI scans were normalized because its data provides relative values. As this is non-trivial, a generalized intensity-scale method was employed based on the work by Madabhushi and Udupa (28). This method performs piecewise-linear normalization utilizing tissue-specific landmarks. These landmarks included the brainstem and the stereotactic G-frame fiducial markers (Elekta AB, Stockholm, Sweden) in the T1- and T1CE-scans. Additionally, for T2-scans, the cerebrospinal fluid was included as landmark.

Third, after normalization, the following radiomic features were extracted.

1) Twenty first-order statistics (FOS) features: statistical properties (e.g., average and variance) of all voxel values, ignoring spatial interaction between image voxels.
2) Four Minkowski functionals (MFs) (29,30): morphological properties from groups of voxels whose intensity is above a specific threshold.
3) Four gray-level co-occurrence matrix (GLCM) features (31): spatial distribution properties of gray-levels in the image voxels, depending on inter-voxel distances, viewing angles, and their quantization levels.
4) Thirteen gray-level size zone matrix (GLSZM) features (32): statistical properties on the size of homogeneously enhanced zones for each gray-level, depending on their quantization levels.

The illustrations in Figure 3 demonstrate a graphical explanation of the MFs, while Figure 4 depicts explanations of the GLCM and GLSZM calculations. All features are calculated on the complete tumor, thus in three-dimensional space.

Fourth, the extracted features are applied to an sML stage to train predictive binary classification models, classifying either true tumor progression or long-term tumor control. Training was performed on a single feature vector, i.e., an array of numerical values, for each individual feature extractor. This is to prevent overfitting, caused by too many included features. If the number of features is large compared with the number of tumors, each tumor can be uniquely identified by one individual feature vector. This results in high prediction scores, but the resulting model may have unreliable performance on new data. A single feature vector is individually created per MRI modality. Furthermore, for the combination of all three MRIs, the individual feature vectors are combined. Together with all the different parameter settings for each feature extractor and the

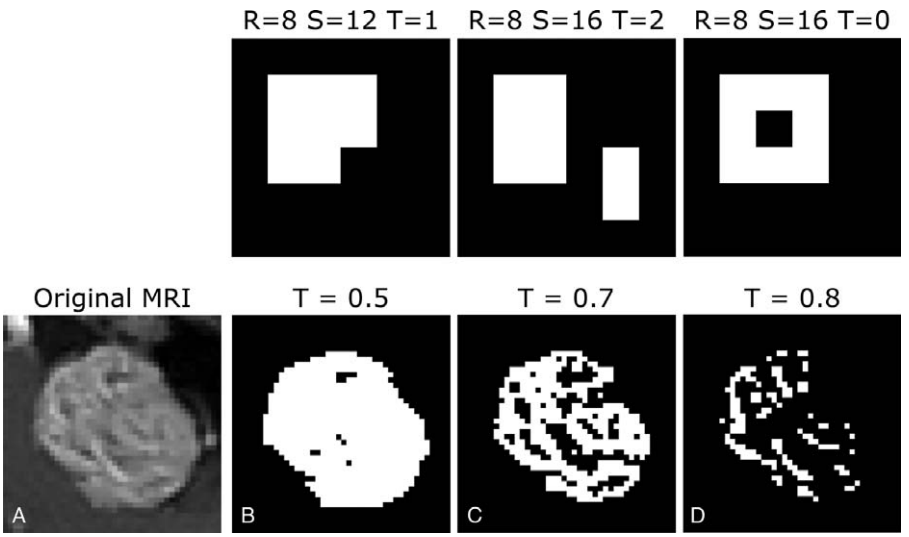

FIG. 3. Visualization of the Minkowski functionals (MFs). The top row represents the visualization of the calculations of the MFs in two-dimensional images. In these simple images of 6 × 6 pixels, we can calculate the number of white pixels (*R*), the length of the boundary of all white shapes (*S*), and the number of white shapes minus the number of black shapes within white shapes (*T*). As such, each image is represented by three values. For three-dimensional images like MRIs, four metrics can be calculated in the same way: 1) the number of white voxels, 2) the number of open faces, 3) the number of open edges, and 4) the number of open vertices. This can be done for multiple threshold levels, resulting in different binary representations of the tumor. An example of these representations can be found at the bottom row of the figure. Here, part A represents the original MRI image. Parts B, C, and D were generated with thresholds *T* equal to 0.5, 0.7, and 0.8, respectively. MRI indicates magnetic resonance imaging.

different settings for the employed sML algorithm, a large number of models are trained and evaluated. All parameters, settings, and calculated features per feature extractor are listed in Table 1 in the Supplemental Digital Content file SDC-Tables, http://links.lww.com/MAO/B84. For implementation of support vector machines, training was carried out by the classification learner application from MATLAB (version 2018b, the MathWorks Inc., Natick, MA).

Fifth, with increasing tumor volumes the number of tumor voxels increases, thereby expanding the amount of texture information. Therefore, we explored the impact of the tumor volume on the prediction results, by imposing various volume thresholds for specific selections of the data. The evaluated volume thresholds were 2, 3, 4, and 5 cm^3.

Finally, validation of the resulting individual models was performed using 10-fold cross-validation. In this method, depicted in Figure 1, the model was subsequently trained with 90% of the MRIs and validated on the 10% left-out scans. This was repeated 10 times, each time leaving out a different set of MRIs. This method resulted in values for accuracy (ACC; correct prediction rate), sensitivity (SENS; proportion of actual controlled tumors correctly identified), and specificity (SPEC; proportion of actual progressed tumors correctly identified). Furthermore, bootstrapping was performed to determine the area under the receiver operating characteristic (AUC; degree of distinction between the two prediction classes), including confidence bounds. An AUC value equal to 1 translates to the ability to perfectly distinguish the two classes, while a value of 0.5 can be interpreted as random selection.

RESULTS

Cohort

After exclusion of patients with NF2, previous treatment, and small tumors, 379 patients were extracted from a database of 735 patients. Of these, 30 patients (7.9%) did not have a post-SRS follow-up beyond 2 years, either due to salvage treatment for their VS (seven patients, 1.8%), or because they were lost to follow-up (23 patients, 6.1%). After excluding these, a total of 349 patients with a median follow-up of 74 months were identified for this study. Of these 349 patients, 30 (8.6%) needed salvage treatment due to recurrent tumor progression. Of the remaining patients, 13 (3.7%) displayed a volumetric tumor progression. By combining these two groups, a so-called progressed tumor cohort of 43 patients (12%) was defined. A total of 42 patients were identified as having obtained long-term tumor control, according to our strict definition of at least 129 months absence of true tumor progression. As such, a total of 85 patients were included in the model training.

Evaluation of Radiomic Features

After stratifying for volume, the total number of patients per volume threshold per group can be found in Table 1. The results of the optimal model for each individual feature extractor, including the additional volume thresholding, can be found in Table 2. The highlights of the obtained results are described below.

The optimal model trained with FOS features, without any additional volumetric threshold, obtained ACC, SENS, SPEC, and AUC values of 0.67, 0.74, 0.60, and 1.00, respectively. The lower- and upper bounds for the AUC value were 0.99 and 1.00, respectively. For increasing volume thresholds, the results increased up to 0.83, 0.94, 0.64, and 0.88 for tumors larger than 5 cm^3, with lower- and upper bounds of the AUC value of 0.59 and 0.96, respectively. Even though the ACC, SENS, and SPEC increased, the AUC value decreased, with a lower

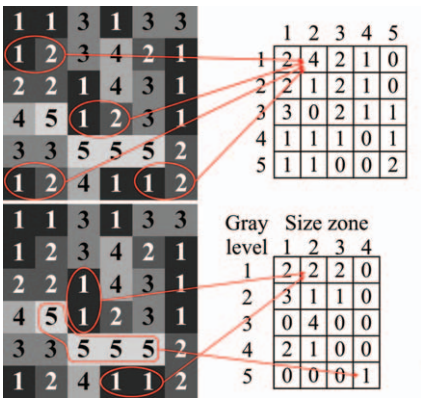

FIG. 4. Graphical representation of the calculations of (*top*) the gray-level co-occurrence matrices (GLCMs), and (*bottom*) the gray-level size-zone matrices (GLSZMs) using the same simple gray-level image. The GLCM is calculated by counting inter-pixel relations. These relations depend on 1) the inter-pixel distance (in this example equal to unity), 2) the inter-pixel angle (in this illustration 0 degrees), and 3) the number of quantization levels, which is the number of different pixel values (here equal to 5). The resulting GLCM matrix (*top-right*) is calculated by counting the number of each specific combination of pixel-pairs. In this example, the pair "1–2" is highlighted. The resulting value in the corresponding position of the final matrix is equal to "4," as there are four pairs "1–2." For the GLSZM (bottom), the number of zones with equal-valued connected pixels and specific size is counted. The pixel values depend on the number of quantization levels, and for different levels distinct GLSZMs can be calculated. In this figure, some example zones are highlighted. These zones consist of connected pixels with the values "1" and "5" of size 2 and 4, respectively. As there are two zones with pixel value "1" of size 2, the corresponding position in the resulting matrix becomes "2." The same can be done for the zone with the value "5." This results in a "1" on the corresponding position in the matrix. From each of these matrices a single feature vector is calculated incorporating the above-described statistics, which is then employed for training.

bound only slightly above 0.5. This is most likely caused by the reduced number of tumors in the dataset, resulting in a less robust model.

The performance scores of the optimal MF-based model obtained ACC, SENS, SPEC, and AUC values of 0.68, 0.64, 0.73, and 0.96, respectively. The lower- and upper bounds for the AUC value were 0.90 and 0.98, respectively. For increasing volume thresholds, the results are increasing up to 0.76, 0.83, 0.64, and 0.88 for tumors larger than 5 cm^3, with AUC lower- and upper bounds of 0.66 and 0.96. Again, the AUC value decreased. Furthermore, the specificity decreases significantly as well. However, the lower bound of the AUC value now remains well above the 0.5 value.

TABLE 1. *Number of patients in both cohorts after additional volume thresholding*

Volume Threshold (cm^3)	True Tumor Progression	Long-term Tumor Control
–	42	43
2	38	38
3	30	28
4	22	18
5	18	11

TABLE 2. *Classification results obtained using radiomics features and additional volume thresholding (in cm^3)*

Feature Extractor	Prediction Score	–	2 cm^3	3 cm^3	4 cm^3	5 cm^3
FOS	ACC	0.67	0.63	0.73	0.66	0.83
	SENS	0.74	0.74	0.80	0.68	0.94
	SPEC	0.60	0.53	0.64	0.65	0.64
	AUC	1.00	0.86	1.00	0.94	0.88
MF	ACC	0.68	0.71	0.67	0.72	0.76
	SENS	0.64	0.74	0.73	0.64	0.83
	SPEC	0.73	0.69	0.60	0.82	0.64
	AUC	0.96	0.91	0.97	0.76	0.88
GLCM	ACC	0.77	0.75	0.76	0.82	0.83
	SENS	0.71	0.76	0.80	0.86	0.83
	SPEC	0.83	0.73	0.72	0.76	0.82
	AUC	0.93	0.92	0.87	0.90	0.99
GLSZM	ACC	0.71	0.73	0.67	0.77	0.69
	SENS	0.79	0.71	0.80	0.86	0.78
	SPEC	0.63	0.74	0.52	0.65	0.55
	AUC	0.88	0.78	0.99	0.84	1.00

The different feature extractors are first-order statistics (FOS), Minkowski Functionals (MFs), gray-level co-occurrence matrices (GLCMs), and gray-level size-zone matrix (GLSZMs). The results are given in accuracy (ACC), sensitivity (SENS), specificity (SPEC), and area under the receiver operating characteristic (AUC).

For the GLSZM-trained models, the optimal model obtained ACC, SENS, SPEC, and AUC values of 0.71, 0.79, 0.63, and 0.88, respectively. The lower- and upper bounds for the AUC value were 0.75 and 0.94, respectively. For the additional volume thresholds, these results remain comparable, depending on the specific threshold. The best results are obtained for tumors larger than 4 cm^3, having an ACC, SENS, SPEC, and AUC of 0.77, 0.86, 0.65, and 0.84, respectively, with AUC lower- and upper bounds of 0.67 and 0.95, respectively.

For the GLCM-based features the ACC, SENS, SPEC, and AUC values were 0.77, 0.71, 0.83, and 0.93, respectively. The lower- and upper bounds of the AUC value were 0.83 and 0.98, respectively. If additional volume thresholding was applied, these results increased to 0.83, 0.83, 0.82, and 0.99, respectively, for tumors larger than 5 cm^3. With this threshold, the lower- and upper bound of the AUC value were 0.94 and 1.00, respectively. In contrast with the previous feature extractors, the AUC values of GLCM features improved for increasing tumor volumes, including the confidence bounds.

DISCUSSION

This work concentrated on finding indicators for prediction of the long-term treatment response of SRS-treated VS patients. The ability to a priori predict such a treatment response can significantly impact the

treatment selection process and may improve the overall treatment outcome. Recently, there is an increasing interest of radiomics in various oncological fields, including intracranial tumors, in relation to their specific pathology and treatment response (33–37). For VS tumors, only incidental literature is available and concerns the prediction of early treatment response, i.e., radiation-induced swelling (38,39). The present paper is, to the best of our knowledge, the first research to focus on the prediction of long-term tumor control following SRS treatment of VS tumors, using MRI-based radiomics. We evaluated tumor-specific parameters obtained from conventional MRI scans in a large database with long follow-up, enabling a high-throughput mining of MRI data which can then be exploited for a machine learning approach. Furthermore, we employed objective definitions for true tumor progression and long-term tumor control. The highest obtained prediction accuracy, sensitivity, specificity, and AUC values were 0.83, 0.83, 0.82, and 0.99, respectively. In other words, the best-performing model is able to correctly predict the treatment outcome in our own data in 83% of all cases (accuracy), progressed tumor in 82% of the cases (specificity), and controlled tumor in 83% of the cases (sensitivity). The most-predictive features were based on GLCMs and in tumors larger than 5 cm^3. These GLCMs measure the distribution of co-occurring voxels at a given gray-value, revealing certain properties about the spatial distribution of the gray-levels in the image. Furthermore, the obtained results show that the prediction improves for increasing tumor volumes. This is most likely caused by the amount of voxels: larger tumors have more tumor voxels, resulting in an increased amount of texture information. However, for increasing volume thresholds, the number of available tumors decreases. This can cause overfitting in the machine learning stage, which results in models that are too fine-tuned to the data, giving high prediction results on the training data. As such, these models may not be robust in predicting the treatment outcome on new unseen data.

A problematic aspect in predicting treatment outcome is the classification of long-term tumor control and true tumor progression, which is crucial in a supervised machine learning approach. Generally, treatments are classified as failed if salvage treatment is needed. Although this is a valid clinical definition, various centers use different motivations before considering salvage treatment (27). As such, there is no clear consensus on whether a treatment has failed. Furthermore, subtle progressions may have been missed in the clinical setting, using linear measurements. This is why we implemented an objective measure for failure, i.e., true tumor progression, using the volumetric tumor response. Moreover, it is hypothesized that radiation-induced swelling is related to different radiobiological aspects of the tumor compared with true tumor progression. To avoid the misclassification of this phenomenon, true tumor progression was defined based on volumetric tumor assessments beyond 2 years after treatment. Long-term tumor control is also difficult to define, as tumor progression can occur many years after SRS. Kondziolka et al. (40) reported that after 4 years following SRS, no further increase in volumes were identified. Contrarily, Hasegawa et al. (41) reported that only after 10 years of follow-up, no more treatment failures occurred. In our extensive database, we determined the latest-occurring tumor progression at 129 months following treatment. Therefore, we adopted this as cutoff for determining tumor control. We realize that these definitions may have had an impact on the obtained results, e.g., tumors now classified as controlled tumor may show tumor progression in the future. However, it can be argued that tumor progression is a very unlikely phenomenon beyond 10 years after treatment.

Another possible confounder in this study is its retrospective character. As a consequence of our long-term tumor control definition, the implemented algorithm used MRIs that were at least 10 years old to find aspects of tumor texture that distinguish tumor control from progression. However, MRIs have improved in the course of time. It is therefore possible that the currently used conventional MRIs exhibit more detailed radiomic features, leading to improved SRS outcome prediction.

Finally, the obtained results are based on the data from a single institution. As such, the obtained models may suffer from input bias errors. Furthermore, overfitting of the trained models is a concern for larger tumors, as the number of tumors that meet our strict criteria of long-term tumor control and true tumor progression, decreases significantly with increasing volume thresholds. To prove the robustness of the results, the obtained algorithms need to be validated on large datasets from multiple centers with sufficiently long follow-up.

Nevertheless, the results of this study show that prediction of long-term tumor control after SRS treatment of larger VS tumors is feasible with the use of radiomic features. This radiomics-based information can potentially be used in a clinical decision support system. Individual MRI scans can serve as an input to software, which contains a well-trained tumor texture model. Such a system can then present the patient and treating physician with prediction scores, facilitating the selection of a personalized optimal treatment strategy.

REFERENCES

1. Lin XEP, Crane XBT, Lin EP. The management and imaging of vestibular schwannomas. *AJNR Am J Neuroradiol* 2017;38:2034–43.
2. Wolbers JG, Dallenga AH, Romero AM, et al. What intervention is best practice for vestibular schwannomas? A systematic review of controlled studies. *BMJ Open* 2013;3:e001345.
3. Kondziolka D, Mousavi SH, Kano H, et al. The newly diagnosed vestibular schwannoma: radiosurgery, resection, or observation? *Neurosurg Focus* 2012;33:E8.
4. Abou-Al-Shaar H, Azab MA, Karsy M, et al. Assessment of costs in open surgery and stereotactic radiosurgery for vestibular schwannomas. *J Neurosurg* 2018;131:561–8.
5. Schnurman Z, Golfinos JG, Epstein D, et al. Comparing costs of microsurgical resection and stereotactic radiosurgery for vestibular schwannoma. *J Neurosurg* 2019;131:1347–682.
6. van de Langenberg R, Hanssens PEJ, van Overbeeke JJ, et al. Management of large vestibular schwannoma. Part I. Planned subtotal resection followed by Gamma Knife surgery: radiological and clinical aspects. *J Neurosurg* 2011;115:875–84.

7. Chung WY, Pan DHC, Lee CC, et al. Large vestibular schwannomas treated by Gamma Knife surgery: long-term outcomes. *J Neurosurg* 2010;113 (suppl):112–21.
8. Lefranc M, Da Roz LM, Balossier A, et al. Place of gamma knife stereotactic radiosurgery in grade 4 vestibular schwannoma based on case series of 86 patients with long-term follow-up. *World Neurosurg* 2018;114:e1192–8.
9. Huang C-W, Tu H-T, Chuang C-Y, et al. Gamma Knife radiosurgery for large vestibular schwannomas greater than 3 cm in diameter. *J Neurosurg* 2018;128:1380–7.
10. Iorio-Morin C, Alsubaie F, Mathieu D. Safety and efficacy of gamma knife radiosurgery for the management of Koos Grade 4 vestibular schwannomas. *Neurosurgery* 2016;78:521–30.
11. Milligan BD, Pollock BE, Foote RL, et al. Long-term tumor control and cranial nerve outcomes following Gamma knife surgery for larger-volume vestibular schwannomas: clinical article. *J Neurosurg* 2012;116:598–604.
12. Bailo M, Boari N, Franzin A, et al. Gamma knife radiosurgery as primary treatment for large vestibular schwannomas: clinical results at long-term follow-up in a series of 59 patients. *World Neurosurg* 2016;95:487–501.
13. Van De Langenberg R, Hanssens PEJ, Verheul JB, et al. Management of large vestibular schwannoma. Part II. Primary Gamma Knife surgery: radiological and clinical aspects - clinical article. *J Neurosurg* 2011;115:885–93.
14. Zeiler FA, Bigder M, Kaufmann A, et al. Gamma knife radiosurgery for large vestibular schwannomas: a Canadian experience. *Can J Neurol Sci* 2013;40:342–7.
15. Wiet RJ, Mamikoglu B, Odom L, et al. Long-term results of the first 500 cases of acoustic neuroma surgery. *Otolaryngol Head Neck Surg* 2001;124:645–51.
16. Anderson DE, Leonetti J, Wind JJ, et al. Resection of large vestibular schwannomas: facial nerve preservation in the context of surgical approach and patient-assessed outcome. *J Neurosurg* 2005;102:643–9.
17. Jung S, Kang SS, Kim TS, et al. Current surgical results of retrosigmoid approach in extralarge vestibular schwannomas. *Surg Neurol* 2000;53:370–8.
18. Lee HJ, Kim MJ, Koh SH, et al. Comparing outcomes following salvage microsurgery in vestibular schwannoma patients failing gamma-knife radiosurgery or microsurgery. *Otol Neurotol* 2017;38:1339–44.
19. Charabi S, Mantoni M, Tos M, et al. Cystic vestibular schwannomas: neuroimaging and growth rate. *J Laryngol Otol* 1994;108:375–9.
20. Spickler E, Lufkin R, Teresi L, et al. MR of hemorrhagic acoustic neuromas. *Comput Med Imaging Graph* 1991;15:333–7.
21. Gomez-Brouchet A, Delisle MB, Cognard C, et al. Vestibular schwannomas: correlations between magnetic resonance imaging and histopathologic appearance. *Otol Neurotol* 2001;22:79–86.
22. de Vries M, Hogendoorn PCW, Briaire-de Bruyn I, et al. Intratumoral hemorrhage, vessel density, and the inflammatory reaction contribute to volume increase of sporadic vestibular schwannomas. *Virchows Arch* 2012;460:629–36.
23. Mindermann T, Schlegel I. Grading of vestibular schwannomas and corresponding tumor volumes: ramifications for radiosurgery. *Acta Neurochir (Wien)* 2013;155:71–4.
24. Konstantina Kouro TP, Exarchos KP, Exarchos MV, et al. Machine learning applications in cancer prognosis and prediction. *Comput Struct Biotechnol J* 2015;13:8–17.
25. Hayhurst C, Zadeh G. Tumor pseudoprogression following radiosurgery for vestibular schwannoma. *Neuro Oncol* 2012;14: 87–92.
26. van de Langenberg R, Dohmen AJC, de Bondt BJ, et al. Volume changes after stereotactic LINAC radiotherapy in vestibular schwannoma: control rate and growth patterns. *Int J Radiat Oncol Biol Phys* 2012;84:343–9.
27. Langenhuizen PPJH, Zinger S, Hanssens PEJ, et al. Influence of pretreatment growth rate on Gamma Knife treatment response for vestibular schwannoma: a volumetric analysis. *J Neurosurg* 2019;131:1347–682.
28. Madabhushi A, Udupa JK. New methods of MR image intensity standardization via generalized scale. *Med Phys* 2006;33:3426–34.
29. Hadwiger H. *Lectures on content, surface, and isoperimetry [title translated from German]*. Verlin, Heidelberg: Springer Berlin Heidelberg; 1957.
30. Li X, Mendonça PRS, Bhotika R. Texture analysis using Minkowski functionals.*Proc SPIE 8314, Medical Imaging 2012: Image Processing*. 2012. 83144Y.
31. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. *IEEE Trans Syst Man Cybern* 1973;SMC-3:610–21.
32. Thibault G, Fertil B, Navarro C, et al. Texture indexes and gray level size zone matrix application to cell nuclei classification. *Pattern Recognit Inf Process* 2009;140–5.
33. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology* 2016;278:563–77.
34. Peng L, Parekh V, Huang P, et al. Distinguishing true progression from radionecrosis after stereotactic radiation therapy for brain metastases with machine learning and radiomics. *Int J Radiat Oncol Biol Phys* 2018;102:1236–43.
35. Tiwari P, Prasanna P, Wolansky L, et al. Computer-extracted texture features to distinguish cerebral radionecrosis from recurrent brain tumors on multiparametric MRI: a feasibility study. *Am J Neuroradiol* 2016;37:2231–6.
36. Zhang Z, Yang J, Ho A, et al. A predictive model for distinguishing radiation necrosis from tumour progression after gamma knife radiosurgery based on radiomic features from MR images. *Eur Radiol* 2018;28:2255–63.
37. Yang D, Rao G, Martinez J, et al. Evaluation of tumor-derived MRI-texture features for discrimination of molecular subtypes and prediction of 12-month survival status in glioblastoma. *Med Phys* 2015;42:6725–35.
38. Speckter H, Santana J, Bido J, et al. Texture analysis of standard magnetic resonance images to predict response to gamma knife radiosurgery in vestibular schwannomas. *World Neurosurg* 2019;132:e228–34.
39. Langenhuizen P, Sebregts S, Zinger S, et al. Prediction of transient tumor enlargement using MRI tumor texture after radiosurgery on vestibular schwannoma. *Med Phys* 2020;47:1692–701. Advance online publication.
40. Kondziolka D, Lunsford LD, Mclaughlin MR, et al. Long-term outcomes after radiosurgery for acoustic neuromas. *N Engl J Med* 1998;339:1426–33.
41. Hasegawa T, Kida Y, Kato T, et al. Long-term safety and efficacy of stereotactic radiosurgery for vestibular schwannomas: evaluation of 440 patients more than 10 years after treatment with Gamma Knife surgery. *J Neurosurg* 2013;118:557–65.