¹⁸F-FDOPA PET for the non-invasive prediction of glioma molecular parameters: a radiomics study

Timothée Zaragori^{1,2}; Julien Oster²; Véronique Roch¹; Gabriela Hossu^{2,3}; Mohammad B. Chawki¹; Rachel Grignon¹; Celso Pouget^{4,5}; Guillaume Gauchotte^{4,5}; Fabien Rech^{6,7}; Marie Blonski^{7,8}; Luc Taillandier^{7,8}; Laëtitia Imbert^{1,2}; Antoine Verger^{1,2}

¹*Université de Lorraine, CHRU-Nancy, Department of Nuclear Medicine & Nancyclotep Imaging platform, F-54000, Nancy, France;*

²*Université de Lorraine, INSERM, IADI, UMR 1254, F-54000, Nancy, France;*

³*CIC 1433 Innovation Technologique, INSERM, Université de Lorraine, CHRU Nancy, F-54000, Nancy, France;*

⁴*Université de Lorraine, CHRU-Nancy, Department of Pathology, F-54000, Nancy, France;*

⁵*Université de Lorraine, INSERM U1256, F-54000, Nancy, France;*

⁶*Université de Lorraine, CHRU-Nancy, Department of Neurosurgery, F-54000, Nancy, France;*

⁷*Université de Lorraine, Centre de Recherche en Automatique de Nancy CRAN, CNRS UMR 7039, F-54000, Nancy, France;*

⁸*Université de Lorraine, CHRU-Nancy, Department of Neuro-oncology, F-54000, Nancy, France.*

Address for correspondence:

Pr Antoine Verger, M.D., Ph.D.

Médecine Nucléaire, Hôpital de Brabois, CHRU-Nancy, Allée du Morvan, 54500 Vandoeuvre-les-Nancy, France.

Tel: (+33)3.83.15.39.09; Fax: (+33).3.83.15.38.39; E-mail: a.verger@chru-nancy.fr

First author:

Timothée Zaragori.

Médecine Nucléaire, Hôpital de Brabois, CHRU-Nancy, Allée du Morvan, 54500 Vandoeuvre-les-Nancy, France.

Tel: (+33)3.83.15.39.09; Fax: (+33).3.83.15.38.39; E-mail: timothee.zaragori@univ-lorraine.fr

Word count: 5787

Running title: ^{18}F -FDOPA radiomics in gliomas

ABSTRACT

Purpose: The assessment of gliomas by ^{18}F -FDOPA PET imaging in adjunct to MRI showed high performance by combining static and dynamic features to non-invasively predict the isocitrate dehydrogenase (IDH) mutations and the 1p/19q co-deletion, which the World Health Organization classified as significant parameters in 2016. The current study evaluates whether other ^{18}F -FDOPA PET radiomics features further improve performance and the contributions of each of these features to performance.

Methods: Our study included seventy-two, retrospectively selected, newly diagnosed, glioma patients with ^{18}F -FDOPA PET dynamic acquisitions. A set of 114 features, including conventional static features and dynamic features as well as other radiomics features were extracted and machine-learning models trained to predict IDH mutations and the 1p/19q co-deletion. Models were based on a machine-learning algorithm built from stable, relevant, and uncorrelated features selected by hierarchical clustering followed by a bootstrapped feature selection process. Models were assessed by comparing area under the curve (AUC) using a nested cross-validation approach. Feature importance was assessed using SHapley Additive exPlanations (SHAP) values.

Results: The best models were able to predict IDH mutations (logistic regression with L2 regularization) and the 1p/19q co-deletion (support vector machine with radial basis function kernel) with an AUC of 0.831[0.790;0.873] and 0.724[0.669;0.782] respectively. For the prediction of IDH mutations, dynamic features were the most important features in the model (TTP: 35.5%). In contrast, other radiomics features were the most useful for predicting the 1p/19q co-deletion (up to 14.5% of importance for the small zone low grey level emphasis).

Conclusions: ^{18}F -FDOPA PET is an effective tool for the non-invasive prediction of glioma molecular parameters using a full set of amino-acid PET radiomics features. The contribution of each feature set shows the importance of systematically integrating dynamic acquisition for the prediction of the IDH mutations as well as developing the use of radiomics features in routine practice for the prediction the 1p/19q co-deletion.

Keywords: ^{18}F -FDOPA PET; radiomics; glioma; machine learning; WHO 2016 classification

INTRODUCTION

The World Health Organization (WHO) 2016 classification of gliomas shifted from a purely histology based approach (1) to the integration of molecular parameters (2). These parameters notably include isocitrate dehydrogenase (IDH) mutations and the 1p/19q co-deletion which are determinant in the patient's prognosis (3).

PET imaging using radiolabeled amino-acids is currently recommended as an adjunct to MRI by the Response Assessment in Neuro-Oncology group (4), with the ability of non-invasively predicting molecular parameters using among others the 6-¹⁸F-fluoro-L-DOPA (¹⁸F-FDOPA) (5). One advantage of such an image-derived non-invasive prediction is the ability of studying the whole tumor area metabolism, while histopathological analyses only consider part of the tumor. In a recent series of 58 newly diagnosed gliomas, ¹⁸F-FDOPA PET imaging was found to predict IDH mutational status with an accuracy of 74% thanks to the integration of dynamic features (5).

This correlated with the emergence of the radiomics field which extracts numerous quantitative features from images, based on classical static and dynamic features (5) but also other features such as morphological and textural features (6). These radiomics studies helped to non-invasively predict specific tumor phenotypes by identifying advanced imaging biomarkers representative of the intra-tumoral heterogeneity with machine learning algorithms (7).

Radiomics is increasingly being used in the PET characterization of gliomas at the initial diagnosis (8), even though only a minority of studies to date have been performed with the amino-acid tracers recommended for glioma assessment (9–11). It has therefore become paramount to understand the potential significance of particular imaging features i.e. static, dynamic and other radiomics features to predict IDH mutations as well as the 1p/19q co-deletion. To the best of our

knowledge, no study has to date assessed the potential benefits of amino-acid PET imaging in the age of radiomics, to characterize the molecular parameters of gliomas according to the WHO 2016 classification, by exploiting amino-acid PET acquisition as a whole.

The current study therefore aims to determine to what extent the high-throughput extraction of advanced imaging biomarkers can predict these molecular parameters and the contributions of each of these ^{18}F -FDOPA PET features to performance.

MATERIALS AND METHODS

Patients

We retrospectively selected newly diagnosed glioma patients that were investigated by ^{18}F -FDOPA PET for suspected glioma components in the Department of Nuclear Medicine at the CHRU of Nancy, between November 2012 and November 2019. Only included patients satisfying the following selection criteria were included: (i) patients with a neuropathological diagnosis of grade II, III and IV gliomas according to the WHO 2016 classification (2) with a maximum delay from the time of the ^{18}F -FDOPA PET of 150 days for diffuse grade II or III gliomas and 60 days for glioblastomas in line with previous published time ranges (5,12,13), (ii) patients with a dynamic analysis and available raw data and (iii) patients with a visually abnormal ^{18}F -FDOPA uptake, i.e. by excluding isometabolic and photopenic gliomas (14,15). A flowchart of patient selection is shown in Figure 1. The institutional ethics committee (Comité d’Ethique du CHRU de Nancy - FRANCE) approved the evaluation of retrospective patient data on August 26, 2020. The trial was registered at ClinicalTrials.gov (NCT04469244). This research complied with the principles of the Declaration of Helsinki. Informed consent was obtained from all individuals included in the study.

^{18}F -FDOPA PET Acquisition and Image Reconstruction

Patients were instructed to fast for at least 4 hours prior to the examination, and some patients were administered Carbidopa one hour prior to the examination to increase the uptake of the tracer in the brain (16). Patients were scanned on a conventional Biograph PET/computed tomography (CT) camera (Biograph 6 True Point, Siemens Healthineers, Erlangen, Germany) and on a digital Vereos PET/CT camera (Vereos, Philips, Cleveland, Ohio). A CT scan was first recorded for each patient, immediately followed by a 30 min 3D list-mode PET recording initiated during the bolus injection of 3 MBq (conventional) or 2 MBq (digital) of ^{18}F -FDOPA per kilogram

of body weight. Dynamic PET images consisted of 30, one minute frames each, while static PET images were reconstructed from the list mode data acquired 10 to 30 min post-injection (10,17) (Figure 2.1).

The conventional PET/CT static and dynamic images were reconstructed using the OSEM 2D algorithm (2 iterations, 21 subsets, 4-mm Gaussian post-reconstruction filter) without time-of-flight information and 256 x 256 x 148 voxels of 2.7 x 2.7 x 3 mm³ with a slice spacing of 1.5 mm for the static and dynamic images. The digital static images were reconstructed using the time-of-flight information and a high-resolution protocol with the OSEM 3D algorithm (2 iterations, 10 subsets, a deconvolution of the point spread function (PSF)) and 256 x 256 x 164 voxels of 1 x 1 x 1 mm³ while a protocol with a lower spatial resolution was used for the dynamic images in order to limit the level of noise, i.e. 3 iterations, 15 subsets, without PSF and 128 x 128 x 82 voxels of 2 x 2 x 2 mm³ (18).

All images were corrected for attenuation using CT, dead time, random and scattered coincidences during the reconstruction process.

Segmentation

Volumes of interest for tumor, contralateral healthy brain and striatum were defined on the static image using the LifeX software (lifexsoft.org) (19).

This healthy brain uptake was measured in crescent-shaped volume of interest (VOI) manually positioned on the unaffected hemisphere so as to comprise both white and grey matter, as recommended (10).

Tumor VOI was segmented semi-automatically using a threshold of 1.6 healthy brain SUV_{mean} (10). The striatum was delineated semi-automatically using a threshold of 70% of the SUV_{max} ¹⁸F-FDOPA uptake and corrected manually where required.

Where a VOI was composed of multiple components, the one with the largest volume was retained in the analysis. All final volumes of interest were visually inspected by an experienced physician (A.V.) to ensure the quality of the methods applied. Examples of representative volumes of interest are provided in Figure 2.2.

Image Pre-processing and Extraction of Features

Figure 2.3 illustrates the workflow of image pre-processing and the extraction of features.

Static Features, Mean, maximum and peak tumor-to-normal brain ratios (TBR_{mean} , TBR_{max} , TBR_{peak}) and tumor-to-striatum ratios (TSR_{mean} , TSR_{max} , TSR_{peak}) were respectively computed as SUV_{mean} , SUV_{max} and SUV_{peak} in the tumor VOI divided by the SUV_{mean} in the brain and the striatum VOI. The metabolic tumor volume (MTV) was defined as the volume of the tumor VOI.

Dynamic Features, Each dynamic frame was first registered to the associated CT image, in order to correct for any potential patient movement during the acquisition (20). The brain and tumor time-activity curves (TACs) were extracted by retrieving the SUV_{mean} for each frame respectively in the brain VOI and in the volume corresponding to the SUV_{peak} of the tumor VOI on the static image, to represent the most aggressive part of the tumor (21). To overcome noise effects, a non-linear fit was applied to each TAC (5,22) with a function dedicated to tumoral vascularization (patent WO/2008/053268, entitled "Method and System for Quantification of Tumoral Vascularization") and a non-linear least optimization using Trust Region Reflective algorithm. Given that TBR values are less likely to be influenced by Carbidopa premedication than SUV, TAC_{ratio} representing the evolution of the ratio between tumor and brain fitted TAC were calculated (5). Time-to-peak (TTP) was computed as the delay between the beginning of the dynamic acquisition and the time of the

maximum uptake value and the slope of the linear regression of the data between the 10th and 30th min was then determined (5).

Other Radiomics Features, For patients whose acquisition was performed on the conventional PET camera, isotropic voxel resampling was performed using the SimpleITK Python package following image standard biomarker initiative (ISBI) recommendation to reach a voxel size of 2.7 x 2.7 x 2.7 mm³ with tricubic spline interpolation. Each static image was normalized by the SUV_{mean} in brain VOI in order to compensate for differences in Carbidopa premedication in our population which is known to impact SUV measurements (23). A fixed bin size of 0.1 was used for absolute image discretization in all patients. 105 features were extracted from the tumor VOI including morphological, local intensity, intensity-based statistical, intensity histogram and textural features. A 3D merging strategy was used for textural matrices (24). The extraction of the features was performed based on the guidelines and benchmark values of the IBSI (24) using pyradiomics (<https://github.com/Radiomics/pyradiomics>) or an in-house software for local intensity features that were not available in pyradiomics. The full list of extracted features is given in Supplementary Table 1 and extraction parameters in Supplementary Table 2.

These 114 features were used for the following classification tasks: (i) IDH mutant (IDH+) vs IDH wildtype (IDH-) gliomas and (ii) gliomas with and without 1p/19q co-deletion (1p19q+ vs 1p19q-).

Data Harmonization

All extracted features other than dynamic features were harmonized with the modified ComBat method (25) (<http://github.com/SteinCK/M-ComBat>) using the digital PET device as reference. This harmonization proved to be efficient for removing the device effect without altering

the biological variation (26). Dynamic features were not harmonized since they represent semi-quantitative kinetic features and are not very much affected by the PET device (9).

Model Building

For a given training and validation set all transformations (detailed below) as well as the algorithm were fitted using the training set only. In the validation set, all transformations computed on the training set were applied and the fitted algorithm was used to predict the output.

All static, dynamic, and other features were first normalized with a Z-score normalization. Then, dimensionality reduction was performed by hierarchical clustering (27), where highly correlated features (absolute Spearman correlation coefficient ≥ 0.9) are clustered together. For each cluster, the feature selected was the medoid of the cluster, i.e. the feature that minimizes the dissimilarity with all other features in the cluster. The most relevant and stable uncorrelated features were selected based on 100 bootstrap samples in which the features were ranked using the Wilcoxon score (28) with a global rank computed with the previously published importance score (27). A range of 5 to 15 selected features were tested, in increments of 5.

Finally, selected features were given as an input to a machine learning classification algorithm. Class imbalance for each classification task was addressed by assigning weights inversely related to the class's prevalence in the training data during machine-learning algorithm training. In case of severely imbalanced groups in a given classification task (classification of 1p/19q co-deletion), the Synthetic Minority Over-sampling Technique algorithm (29) was used for oversampling the minority class. The performance of five classifiers implemented were compared: logistic regression with L2 regularization (LR_L2), neural networks (Nnet), random forest (RF), support vector machine with radial basis function kernel (SVM_RBF) and support vector machine with linear kernel (SVM_linear).

Model Performance

Nested cross-validation was used to get an unbiased estimate of the model's performance (30) with repeated stratified cross-validations as inner and outer loops to reduce the influence of the variance on the results (31). Hyperparameters tuning was performed through an inner loop with a 3-fold stratified cross-validation repeated 50 times by optimizing the mean area under the curve (AUC) overall inner folds and using 100 iterations of a random search over the set of hyperparameters described in Supplementary Table 3. The model assessment was performed through the outer loop with a 10-fold stratified cross-validation repeated 5 times by calculating the mean AUC overall outer folds. For this purpose, an in-house software was developed and is freely available online

(https://github.com/TimZaragori/Sklearn_NestedCV/tree/master/Radiomics_gliomas_article).

The whole modeling pipeline is shown in Figure 3. For each classification task, the final model selected was the one with the highest performance.

Statistical Analysis

Categorical variables are expressed as percentages and continuous variables as medians [first quartile; third quartile] because variables did not follow a normal distribution. Intergroup comparisons were performed with the Chi-squared test for categorical variables and the Mann-Whitney test for continuous variables. For each outer fold receiver-operating-characteristic curve, the optimal threshold was computed on the training set by selecting the point on the curve closest to (0,1) and applying it on the validation set to get the predictions. Diagnostic performance (accuracy, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV)) were computed from outer model predictions. *p* values lower than 0.05 were considered

significant. 1000 bootstrap iterations of outer results were used to determine the 95% confidence intervals of each dataset performance. Analyses were performed with the R software version 3.6.2 (R Foundation for Statistical Computing, Vienna, Austria). For each classification, the model introspection was performed using SHapley Additive exPlanations (SHAP) values (32) which provide information on the importance of each feature by taking into account the whole modeling pipeline. SHAP values were computed on each outer model and concatenated to obtain a global and reliable importance score for each feature.

RESULTS

Patients

We retrospectively selected 72 patients (mean age 52.8 [38.4; 64.7] years, 29 women), who underwent a ^{18}F -FDOPA PET for a newly diagnosed glioma, with a histopathological diagnosis either determined from tissue obtained during surgery (n=33) or from biopsies (n=39): 43 IDH+ gliomas from whom 18 1p/19q co-deleted. The time window between PET imaging and neuropathological confirmation was 35 [11.5; 84] days for diffuse grade II or III gliomas and 3 [1; 24] days for glioblastomas. Data for forty-four of the patients (61%) was acquired by conventional PET/CT and twenty-eight (39%) by digital PET/CT. Forty-two (58%) patients were premedicated with Carbidopa. Patients acquired by the conventional PET/CT were already analyzed previously (5). Further details of patients' characteristics are provided in Table 1. The validation of the harmonization and normalization processes are respectively shown in Supplementary Table 4 and Supplementary Table 5.

Predicting IDH Mutations

The combination of LR_L2 and 5 selected features was the best performing model for predicting IDH mutations and yielded an AUC of 0.831 [0.790; 0.873] (details of the predictive performance of this model and the performance of all models for predicting IDH mutations are presented in Table 2 and Supplemental Figure 1, respectively). The TTP was found to be the feature with the highest importance reaching a 35.5% contribution to the overall model's performance (Figure 4). The other dynamic parameter, Slope, was also among the five most contributive features even though with a lower importance (4.3%).

Predicting the 1p/19q Co-Deletion

The 1p/19q co-deletion was best predicted by a model which combined SVM_RBF and 15 features selected (AUC: 0.724 [0.669; 0.782]) (see Supplemental Figure 2 for results of all models for 1p/19q co-deletion and Table 2 for the complete predictive performance of this model). In contrast to the prediction of the IDH mutations, the other radiomics features were the most contributive with the Small zone low grey level emphasis *GLSZM* being the most important feature (14.5%) (Figure 4).

An overview of the impact on the model of each feature according to its values for the interpretation of the radiomics signatures is shown in Figure 5. Representative examples of patients for both classifications with their local interpretation using SHAP values are provided in Figure 6.

DISCUSSION

The current study shows that ^{18}F -FDOPA PET imaging is an efficient tool for the non-invasive prediction of two of the most important molecular parameters of newly diagnosed gliomas according to the WHO 2016 classification, the IDH mutations using dynamic features and the 1p/19q co-deletion using other radiomics features. Using a set of 114 radiomics features, a good performance for predicting these two parameters was achieved, with AUCs of 0.831 and 0.724, respectively. Interestingly, the greatest contribution to the predictive performance for IDH mutations was provided by dynamic features while the contribution of the set of other radiomics features was the most useful for predicting the 1p/19q co-deletion (Table 2, Figures 4 and Figure 5).

Since the emergence of the radiomics field, only a few amino-acid PET studies have been devoted to the characterization of molecular parameters in newly diagnosed gliomas (8). Among these, only two were performed for the two molecular parameters which are key features of the WHO 2016 classification. Lohmann et al. (9) exclusively focused their analysis on predicting IDH genotype while Haubold et al. extended their analysis to various other molecular parameters using MR fingerprints and ^{18}F -FET but did not include the information provided by the dynamic PET acquisitions (11). The objective of our study was to predict the classification of newly diagnosed gliomas according to both IDH mutations and the 1p/19q co-deletion by considering amino-acid PET data as a whole i.e., by extracting a full set of radiomics features including dynamic ones to better characterize the contribution of each group of features (static, dynamic, and other radiomics features) in the prediction of molecular parameters.

For the prediction of IDH mutations, our results are consistent with those of Lohmann et al. using dynamic ^{18}F -FET, in which they reported accuracies of greater than 70% (9) (vs 74.7%

in the present study). Conversely, Haubold et al. reported lower predictive performances for IDH mutations (AUC of 0.639 vs 0.831 in the present study) but without including any dynamic information in their analyses (11). This latter point underlines the fact that dynamic information is crucial for the prediction of IDH mutations, consistent with our previous results, in which dynamic features led to accuracies of predicting IDH mutations of up to 74% (5). For the prediction of the 1p/19q co-deletion, the set of other radiomics features provided the most important contribution (Figures 4 and Figure 5). These statements are supported by the findings of Haubold et al. who reported a much higher prediction of the textural features for the classification of the 1p/19q co-deletion (AUC of 0.858) than IDH mutations (AUC of 0.639). The highly predictive value of dynamic parameters for IDH mutations and other radiomics features for 1p/19q co-deletion is further supported by additional results obtained when comparing the performances of the 114 features dataset to dataset with only static features, static and other radiomics features and, static and dynamic features (please see Supplemental Figure 3).

It is crucial to try and explain glioma signatures in terms of the significance of the selected features (Figures 4 and Figure 5). In the signature of the IDH mutations, IDH wildtype gliomas are characterized by low TTP values. This is consistent with results from our previous study (5) where IDH wildtype gliomas, reputed as more aggressive tumors, are associated with a breakdown of the blood-brain barrier, but also a higher microvessel density and LAT1 expression. All these considerations are consistent with intense wash-in and wash-out phenomena, leading to short TTP times.

The signature interpretation for the 1p/19q co-deletion classification is more complex and can be considered in two parts. Firstly, by focusing on features that have highly negative SHAP values and that help the classifier exclude a potential 1p19q+ glioma. This is the case for high

values of Minimum $1s$, given that $1p19q+$ gliomas exhibit high amino-acid PET uptake (33), as well as low values of Small zone low grey level emphasis $GLSZM$ and high values of Busyness $NGTDM$, since $1p19q+$ gliomas are homogeneous tumors. Secondly, and in contrast to the first interpretation of the signature, features exhibiting positive SHAP values allow the classifier to confirm a potential $1p19q+$ glioma. This is the case for features associated with a non-aggressive behavior such as high values of Sphericity $MORPH$ characterizing spheric tumors or other features from the $GLSZM$ matrix representative of a homogeneous tumor (high values of Large zone emphasis and low values of Zone percentage).

As a limitation, the number of patients is relatively low, particularly the number of glioma patients with $1p/19$ co-deletions, even though it is comparable and even larger compared to other radiomics studies performed on this relatively rare entity (9,11). Our methodology for radiomics modelling (feature selection, nested cross-validation, and adapted corrections for imbalanced datasets) was nevertheless adapted to the number of patients to avoid the risk of overfitting. Nevertheless, validation on an external dataset of the suggested models is still required before transfer to the clinical setting. Another limitation is the use of two different PET cameras for acquisitions which we handled by harmonizing the features of each device using the Modified ComBat method (Supplemental Table 4). In contrast to the ComBat method, the modified method has the advantage of aligning data in the space of a specific device which means that results can be used in routine clinical practice.

CONCLUSIONS

To the best of our knowledge, our current study is the first to focus on predicting the presence of major molecular parameters that are of prognostic interest based on the WHO 2016 glioma classification, by using a full set of amino-acid PET static, dynamic and radiomics features. Interestingly, most of the salient information for the prediction of IDH mutations is provided by the dynamic analysis, which highlights the necessity of systematically including dynamic acquisition as part of routine amino-acid PET imaging. However, for the classification of the 1p/19q co-deletion, the predominant contribution of textural features shows the need to develop the extraction of such features in routine through machine learning models as complementary tools for clinical interpretation.

FUNDING

None

CONFLICT OF INTEREST

No potential conflicts of interest relevant to this article exist.

ACKNOWLEDGEMENTS

The authors thank Petra Neufing for critical review of the manuscript.

KEY POINTS:

Question: To what extent the high-throughput extraction of advanced imaging biomarkers can predict the molecular parameters in gliomas and what is the contribution of each ^{18}F -FDOPA PET features in the performance?

Pertinent findings: ^{18}F -FDOPA PET is an effective tool for the non-invasive prediction of glioma molecular parameters using a full set of amino-acid PET radiomics features, with respective AUC of 0.831 [0.790; 0.873] and 0.724 [0.669; 0.782] for IDH mutations (dynamic features) and the 1p/19q codeletion (radiomics features).

Implications for patient care: The contribution of each feature set shows the importance of systematically integrating dynamic acquisition as well as developing the use of radiomics features in clinical ^{18}F -FDOPA PET imaging.

REFERENCES

1. Louis DN, Ohgaki H, Wiestler OD, et al. The 2007 WHO classification of tumours of the central nervous system. *Acta Neuropathol.* 2007;114:97-109.
2. Louis DN, Perry A, Reifenberger G, et al. The 2016 World Health Organization classification of tumors of the central nervous system: a summary. *Acta Neuropathol.* 2016;131:803-820.
3. Eckel-Passow JE, Lachance DH, Molinaro AM, et al. Glioma groups based on 1p/19q, IDH, and TERT promoter mutations in tumors. *N Engl J Med.* 2015;372:2499-2508.
4. Albert NL, Weller M, Suchorska B, et al. Response Assessment in Neuro-Oncology working group and European Association for Neuro-Oncology recommendations for the clinical use of PET imaging in gliomas. *Neuro-Oncol.* 2016;18:1199-1208.
5. Ginot M, Zaragori T, Marie P-Y, et al. Integration of dynamic parameters in the analysis of 18F-FDopa PET imaging improves the prediction of molecular features of gliomas. *Eur J Nucl Med Mol Imaging.* 2020;47:1381-1390.
6. Visvikis D, Cheze Le Rest C, Jaouen V, Hatt M. Artificial intelligence, machine (deep) learning and radio(geno)mics: definitions and nuclear medicine imaging applications. *Eur J Nucl Med Mol Imaging.* 2019;46:2630-2637.
7. Aerts HJWL, Velazquez ER, Leijenaar RTH, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun.* 2014;5:1-9.
8. Lohmann P, Galldiks N, Kocher M, et al. Radiomics in neuro-oncology: Basics, workflow, and applications. *Methods.* June 2020 [Epub ahead of print].
9. Lohmann P, Lerche C, Bauer EK, et al. Predicting IDH genotype in gliomas using FET PET radiomics. *Sci Rep.* 2018;8:13328.
10. Law I, Albert NL, Arbizu J, et al. Joint EANM/EANO/RANO practice guidelines/SNMMI procedure standards for imaging of gliomas using PET with radiolabelled amino acids and [18F]FDG: version 1.0. *Eur J Nucl Med Mol Imaging.* 2019;46:540-557.
11. Haubold J, Demircioglu A, Gratz M, et al. Non-invasive tumor decoding and phenotyping of cerebral gliomas utilizing multiparametric 18F-FET PET-MRI and MR Fingerprinting. *Eur J Nucl Med Mol Imaging.* 2020;47:1435-1445.
12. Verger A, Stoffels G, Bauer EK, et al. Static and dynamic 18F-FET PET for the characterization of gliomas defined by IDH and 1p/19q status. *Eur J Nucl Med Mol Imaging.* 2018;45:443-451.
13. Fueger BJ, Czernin J, Cloughesy T, et al. Correlation of 6-18F-Fluoro-L-Dopa PET uptake with proliferation and tumor grade in newly diagnosed and recurrent gliomas. *J Nucl Med.* 2010;51:1532-1538.

14. Galldiks N, Unterrainer M, Judov N, et al. Photopenic defects on O-(2-[18F]-fluoroethyl)-L-tyrosine PET: clinical relevance in glioma patients. *Neuro-Oncol.* 2019;21:1331-1338.
15. Zaragori T, Castello A, Guedj E, et al. Photopenic defects in gliomas with amino-acid PET and relative prognostic value: a multicentric 11C-Methionine and 18F-FDOPA PET experience. *Clin Nucl Med.* 2021;46:e36-e37.
16. Schiepers C, Chen W, Cloughesy T, Dahlbom M, Huang S-C. 18F-FDOPA Kinetics in Brain Tumors. *J Nucl Med.* 2007;48:1651-1661.
17. Janvier L, Olivier P, Blonski M, et al. Correlation of SUV-derived indices with tumoral aggressiveness of gliomas in static 18F-FDOPA PET: use in clinical practice. *Clin Nucl Med.* 2015;40:e429-e435.
18. Salvadori J, Imbert L, Perrin M, et al. Head-to-head comparison of image quality between brain 18F-FDG images recorded with a fully digital versus a last-generation analog PET camera. *EJNMMI Res.* 2019;9:61.
19. Nioche C, Orlhac F, Boughdad S, et al. LIFEx: A freeware for radiomic feature calculation in multimodality imaging to accelerate advances in the characterization of tumor heterogeneity. *Cancer Res.* 2018;78:4786-4789.
20. Ye H, Wong K-P, Wardak M, et al. Automated movement correction for dynamic PET/CT Images: evaluation with phantom and patient data. *PLoS ONE.* 2014;9:e103745.
21. Pafundi DH, Laack NN, Youland RS, et al. Biopsy validation of 18F-DOPA PET and biodistribution in gliomas for neurosurgical planning and radiotherapy target delineation: results of a prospective pilot study. *Neuro-Oncol.* 2013;15:1058-1067.
22. Zaragori T, Ginet M, Marie P-Y, et al. Use of static and dynamic [18 F]-F-DOPA PET parameters for detecting patients with glioma recurrence or progression. *EJNMMI Res.* 2020;10:1-10.
23. Hoffman JM, Melega WP, Hawk TC, et al. The effects of carbidopa administration on 6-[18F]fluoro-L-dopa kinetics in positron emission tomography. *J Nucl Med.* 1992;33:1472-1477.
24. Zwanenburg A, Vallières M, Abdalah MA, et al. The Image Biomarker Standardization Initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology.* 2020;295:328-338.
25. Stein CK, Qu P, Epstein J, et al. Removing batch effects from purified plasma cell gene expression microarrays with modified ComBat. *BMC Bioinformatics.* 2015;16:63.
26. Orlhac F, Boughdad S, Philippe C, et al. A postreconstruction harmonization method for multicenter radiomic studies in PET. *J Nucl Med.* 2018;59:1321-1328.

27. Leger S, Zwanenburg A, Pilz K, et al. A comparative study of machine learning methods for time-to-event survival data for radiomics risk modelling. *Sci Rep.* 2017;7:1-11.
28. Parmar C, Grossmann P, Bussink J, Lambin P, Aerts HJWL. Machine learning methods for quantitative radiomic biomarkers. *Sci Rep.* 2015;5:1-11.
29. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic Minority Over-sampling Technique. *J Artif Intell Res.* 2002;16:321-357.
30. Cawley GC, Talbot NLC. On over-fitting in model selection and subsequent selection bias in performance evaluation. *J Mach Learn Res.* 2010;11:2079-2107.
31. Krstajic D, Buturovic LJ, Leahy DE, Thomas S. Cross-validation pitfalls when selecting and assessing regression and classification models. *J Cheminformatics.* 2014;6:10.
32. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. *Adv Neural Inf Process Syst.* 2017;30:4765-4774.
33. Jansen NL, Schwartz C, Graute V, et al. Prediction of oligodendroglial histology and LOH 1p/19q using dynamic [18F]FET-PET imaging in intracranial WHO grade II and III gliomas. *Neuro-Oncol.* 2012;14:1473-1480.

Figure 1. Flowchart of the retrospective selection of study patients.

1

2 **Figure 2.** Detailed workflow of (1.) image acquisition and reconstruction (2.) Volumes of interest

3 segmentation (shown in red contours) and (3.) features extracted where images and volumes of

- 1 interest used are listed in orange and grey rectangles respectively, processing steps in green
- 2 rectangles, intermediate results in blue rectangles and final features in light blue rectangles.
- 3 MTV: metabolic tumor volume; TBR: tumor-to-brain ratio; TSR: tumor-to-striatum ratio; TAC:
- 4 time-activity curve; TTP: time-to-peak.
- 5

Figure 3. Modelling pipeline using a nested cross-validation to obtain an unbiased estimate of model performance. 50 repeats of the 3-fold cross-validation were used as an inner loop and 5 repeats of the 10-fold cross-validation as outer loop. For each fold of outer cross-validation, the hyperparameters of the model were tuned in the inner cross-validation using only outer training data. The best hyperparameters chosen are then used to fit the model to outer training data and the model performance is evaluated on outer test data. Blue and red rectangles respectively denote

1 training and test data for each cross-validation. Green arrows represent the predictions made by
2 the fitted model on the test data and green boxes represent the score calculated from these
3 predictions.

4 OF: Outer Fold; IF: Inner Fold

5

Figure 4. Feature importance derived from outer models for the most contributive features to the prediction of A. the IDH mutations using a logistic regression with L2 regularization and 5 features selected and B. the 1p/19q co-deletion using a support vector machine with radial basis function kernel and 15 features selected. The individual importance of each feature was normalized to the sum of 1.0. In blue, other radiomic features and in orange, dynamic features.

Figure 5. Overview of the impact on the model of each feature according to its values (A. for the prediction of IDH mutations and B. for the prediction of 1p/19q co-deletion). Impact on the model output is shown with SHAP values on the x axis and feature value is displayed in colors (blue and red for respective low and high values). For instance, in part A. for the classification of IDH mutations, high values of TTP (in red) are showing positive SHAP values and thus are associated with the prediction of the positive class (IDH+ gliomas) as opposed to low values of

- 1 TTP (in blue) which exhibit negative SHAP values meaning that they are associated with the
- 2 prediction of the negative class (IDH- gliomas).

3

4

A. IDH wildtype Grade II Astrocytoma

B. IDH mutant and 1p/19q co-deleted Grade III Oligodendroglioma

Figure 6. Representative examples of patients with IDH wildtype glioma (A.) and 1p/19q co-deleted glioma (B.). Top row of each patient shows axial slice of ^{18}F -FDOPA PET (left column), dynamic TBR_{mean} curve (middle column, TTP: light grey dotted line, slope: dark blue dotted line) along with the same slice on a FLAIR MRI (right column). For each patient a graphic representing the contribution of features involved in the model prediction (A. IDH mutations classification, B. 1p/19q co-deletion classification) is presented in bottom row. Red features push the prediction towards the positive class while blue features push towards negative class, the longer the arrow the more impact on the model. Base model values (value of a feature-less model) and final decision values are displayed respectively in italic and bold. For interpretation

1 purposes, feature names along with their value (expressed in Z-score) are displayed under each
2 arrow.

3 TTP: time-to-peak; ZP: zone percentage; DV: difference variance; LZE: large zone emphasis;
4 SZLGE: small zone low grey level emphasis; GLNU: grey level non uniformity; DCNUN:
5 dependence count non uniformity normalized.

6

1 **Table 1.** Patients' characteristics.

n=75	Conventional (n=44)	Digital (n=28)	All (n=72)	p value
Age (years)				0.020*
Median [Range]	42.4 [19.8; 73.7]	57.2 [20.3; 78.6]	51.0 [19.8; 78.6]	
Female gender n (%)	15 (34)	14 (50)	29 (40)	0.273
Primary histopathological type n (%)				0.305
Astrocytoma, IDH-mutant	10 (23)	7 (25)	17 (24)	
Anaplastic Astrocytoma, IDH-mutant	5 (11)	1 (4)	6 (8)	
Astrocytoma, IDH-wildtype	2 (5)	3 (11)	5 (7)	
Anaplastic Astrocytoma, IDH-wildtype	2 (5)	3 (11)	5 (7)	
Oligodendroglioma, IDH-mutant and 1p/19q co-deleted	7 (16)	4 (14)	11 (15)	
Anaplastic Oligodendroglioma, IDH-mutant and 1p/19q co-deleted	7 (16)	0 (0)	7 (10)	
Glioblastoma, IDH-wildtype	10 (23)	9 (32)	19 (26)	
Glioblastoma, IDH-mutant	1 (2)	1 (4)	2 (3)	
Carbidopa premedication n (%)	14 (32)	28 (100)	42 (58)	< 0.001*
CE on MRI n (%)	18 (41)	11 (39)	29 (40)	1

2

3 * $p < 0.05$ for the comparison between conventional PET and digital PET patients

4 CE, contrast enhancement.

Table 2. Mean [95% confidence interval] of model performances for each classification.

IDH+ vs IDH- (LR_L2 algorithm and 5 features selected)					
AUC	Sensitivity (%)	Specificity (%)	Accuracy (%)	PPV (%)	NPV (%)
0.831 [0.790; 0.873]	77.0 [72.5; 81.4]	70.7 [63.0; 78.3]	74.7 [71.4; 78.4]	82.5 [78.5; 86.7]	70.9 [64.9; 76.6]
1p19q+ vs 1p19q- (SVM_RBF algorithm and 15 features selected)					
AUC	Sensitivity (%)	Specificity (%)	Accuracy (%)	PPV (%)	NPV (%)
0.724 [0.669; 0.782]	55.0 [43.0; 66.0]	72.3 [66.7; 78.5]	67.0 [62.5; 71.2]	39.6 [30.4; 48.3]	84.2 [80.5; 88.2]

IDH+: IDH mutant gliomas; *IDH-:* IDH wildtype gliomas; *1p19q+:* 1p/19q co-deleted gliomas; *1p19q-:* gliomas not 1p/19q co-deleted; *LR_L2:* logistic regression with L2 regularization; *SVM_RBF:* support vector machine with radial basis function kernel; *AUC:* area under the receiver operating characteristic curve; *PPV:* positive predictive value; *NPV:* negative predictive value.

Supplementary Table 1. Detail of radiomics features extracted for each dataset according to the IBSI definitions, where applicable. S, D and OR respectively stand for static features, dynamic features and other radiomics features.

TBR: tumor-to-brain ratio; TSR: tumor-to-striatum ratio

S: Static features (in-house software)	
1.	Mean TBR
2.	Maximum TBR
3.	Peak TBR
4.	Mean TSR
5.	Maximum TSR
6.	Peak TSR
7.	Metabolic tumor volume
D: Dynamic TBR features (in-house software)	
1.	Time-to-peak
2.	Slope
OR: Morphological features (pyradiomics)	
1.	Volume (mesh)
2.	Volume (counting)
3.	Surface area (mesh)
4.	Surface to volume ratio
5.	Sphericity
6.	Maximum 3D diameter
7.	Major axis length
8.	Minor axis length
9.	Least axis length
10.	Elongation
11.	Flatness
OR: Local intensity features (in-house software)	
1.	Local intensity peak
2.	Global intensity peak
OR: Intensity-based statistical features (pyradiomics)	
1.	Mean intensity
2.	Intensity variance
3.	Intensity skewness
4.	(Excess) intensity kurtosis
5.	Median intensity
6.	Minimum intensity
7.	10th intensity percentile
8.	90th intensity percentile
9.	Maximum intensity
10.	Intensity interquartile range
11.	Intensity range
12.	Intensity mean absolute deviation
13.	Intensity-based robust mean absolute deviation
14.	Intensity-based energy

15. Root mean square intensity

OR: Intensity histogram features (pyradiomics)

1. Discretized intensity entropy

2. Discretized intensity uniformity

OR: Grey level co-occurrence based features (pyradiomics)

1. Joint maximum

2. Joint average

3. Joint variance

4. Joint entropy

5. Difference average

6. Difference variance

7. Difference entropy

8. Sum average

9. Sum entropy

10. Angular second moment

11. Contrast

12. Dissimilarity

13. Inverse difference

14. Normalized inverse difference

15. Inverse difference moment

16. Normalized inverse difference

17. Inverse variance

18. Correlation

19. Autocorrelation

20. Cluster tendency

21. Cluster shade

22. Cluster prominence

23. Information correlation 1

24. Information correlation 2

OR: Grey level run length based features (pyradiomics)

1. Short runs emphasis

2. Long runs emphasis

3. Low grey level run emphasis

4. High grey level run emphasis

5. Short run low grey level emphasis

6. Short run high grey level emphasis

7. Long run low grey level emphasis

8. Long run high grey level emphasis

9. Grey level non-uniformity

10. Normalized grey level non-uniformity

11. Run length non-uniformity

12. Normalized run length non-uniformity

13. Run percentage

14. Grey level variance

15. Run length variance

16. Run entropy

OR: Grey level size zone based features (pyradiomics)

-
1. Small zone emphasis
 2. Large zone emphasis
 3. Low grey level emphasis
 4. High grey level emphasis
 5. Small zone low grey level emphasis
 6. Small zone high grey level emphasis
 7. Large zone low grey level emphasis
 8. Large zone high grey level emphasis
 9. Grey level non-uniformity
 10. Grey level non-uniformity normalized
 11. Zone size non-uniformity
 12. Zone size non-uniformity normalized
 13. Zone percentage
 14. Grey level variance
 15. Zone size variance
 16. Zone size entropy

OR: Neighborhood grey tone difference based features (pyradiomics)

1. Coarseness
2. Contrast
3. Busyness
4. Complexity
5. Strength

OR: Neighboring grey level dependence based features (pyradiomics)

1. Low dependence emphasis
 2. High dependence emphasis
 3. Low grey level count emphasis
 4. High grey level count emphasis
 5. Low dependence low grey level emphasis
 6. Low dependence high grey level emphasis
 7. High dependence low grey level emphasis
 8. High dependence high grey level emphasis
 9. Grey level non-uniformity
 10. Dependence count non-uniformity
 11. Dependence count non-uniformity normalized
 12. Grey level variance
 13. Dependence count variance
 14. Dependence count entropy
-

Supplementary Table 2. IBSI reporting guidelines

Topic	Item	Description
Patient		
Region of interest	1	Tumor VOI segmented semi-automatically using a threshold of 1.6 SUV _{mean} in healthy brain VOI (crescent-shaped VOI manually positioned on the unaffected hemisphere so as to comprise both white and grey matter)
Patient preparation	2a	Fast for at least 4 hours prior to the examination
	2b	For certain patients: Carbidopa administered one hour prior to the examination to increase uptake of tracer in the brain
	2c	N.A.
Radioactive tracer	3a	¹⁸ F-FDopa
	3b	Bolus
	3c	3 MBq/kg (conventional PET), 2 MBq/kg (digital PET)
	3d	Image acquisition started at the same time as radiotracer injection
	3e	N.A.
Contrast agent	4a	N.A.
	4b	N.A.
	4c	N.A.
	4d	N.A.
	4e	N.A.
Comorbidities	5	N.A.
Acquisition		
Acquisition protocol	6	30 min 3D List-mode acquisition (<i>I</i>)
Scanner type	7	Conventional PET: Biograph 6 True Point, Siemens Healthineers, Erlangen, Germany Digital PET: Vereos, Philips, Cleveland, Ohio
Imaging modality	8	PET
Static/dynamic scans	9a	Dynamic scans
	9b	Dynamic images reconstructed in 30 frames of one minute each
	9c	N.A.
Scanner calibration	10	Checkup, air calibration daily QC (ANSM guidelines)
Patient instructions	11	Not moving
Anatomical motion correction	12	Dedicated head holder
Scan duration	13	30 minutes
Reconstruction		
In-plane resolution	28	Conventional static and dynamic images: 2.7 mm x 2.7 mm Digital static image: 1 mm x 1 mm Digital dynamic images: 2 mm x 2 mm
Image slice thickness	29	Conventional static and dynamic images: 3 mm Digital static image: 1 mm Digital dynamic images: 2 mm
Image slice spacing	30	Conventional static and dynamic images: 1.5 mm Digital static image: 1 mm Digital dynamic images: 2 mm

Reconstruction method	32a	Conventional static and dynamic images: 2D OSEM without time-of-flight information Digital static and dynamic images: 3D OSEM with time-of-flight information
	32b	Conventional static and dynamic images: 2 iterations Digital static image: 2 iterations Digital dynamic images: 3 iterations
	32c	Conventional static and dynamic images: 21 subsets Digital static image: 10 subsets Digital dynamic images: 15 subsets
Point spread function modelling	33	Conventional static and dynamic images: N.A. Digital static image: deconvolution of the point spread function (PSF) with 1 iteration and 6-mm regularization kernel applied on post-reconstruction (vendor defaults) Digital dynamic images: N.A.
Image corrections	34a	CT based attenuation correction
	34b	Dead time, random and scattered coincidences corrections during the reconstruction process
Image registration		
Registration method	37	Static and dynamic images were registered to CT using a 3D rigid registration with mutual information and linear interpolation implemented in SimpleITK.
Image processing – data conversion		
SUV normalization	38	$SUV = \frac{\text{Image values (kBq/mL)}}{\frac{\text{Injected dose (kBq)}}{\text{Patient weight (g)}}}$
Other data conversions	40	N.A.
Image processing – post-acquisition processing		
Anti-aliasing	41	N.A.
Noise suppression	42	N.A.
Post-reconstruction smoothing filter	43	Conventional static and dynamic images: 4 mm Digital static and dynamic images: N.A.
Intensity normalization	46	N.A.
Other post-acquisition processing methods	47	N.A.
Segmentation		
Segmentation method	48a	Semi-automatically
	48b	N.A.
	48c	Tumor VOI segmented semi-automatically using a threshold of 1.6 SUV _{mean} in healthy brain VOI (crescent-shaped VOI manually positioned on the unaffected hemisphere so as to comprise both white and grey matter)
	48d	PET static image
Conversion to mask	49	N.A.
Image processing – image interpolation		
Interpolation method	50a	Tricubic spline interpolation

	50b	Align grid centers
	50c	N.A.
	50d	Set to 0
Voxel dimensions	51	For conventional static image only (on which textural features will be extracted): 2.7 x 2.7 x 2.7 mm ³
Image processing – ROI interpolation		
Interpolation method	53	Linear interpolation
Partially masked voxels	54	Threshold of 0.5
Image processing – re-segmentation		
Re-segmentation methods	55	[0; SUV _{max}]
Image processing – discretization		
Discretization method	56a	Fixed bin size
	56b	FBS = 0.1
	56c	0
Image processing – image transformation		
Image filter	57	N.A.
Image biomarker computation		
Biomarker set	58	See Supplemental Table 1
IBSI compliance	59	IBSI compliant
Robustness	60	N.A.
Software availability	61	For Local intensity features: Home-made python software ¹ For others: pyradiomics v2.2.0b1 ²
Image biomarker computation – texture parameters		
Texture matrix aggregation	62	IS, IH: 3D GLCM, GLRLM: 3D: mrg GLSZM, NGTDM, NGLDM: 3D
Distance weighting	63	No weighting
CM symmetry	64	Symmetric
CM distance	65	1
SZM linkage distance	66	1
DZM linkage distance	67	N.A.
DZM zone distance norm	68	N.A.
NGTDM distance	69	1
NGLDM distance	70	1
NGLDM coarseness	71	0
Machine learning and radiomics analysis		
Diagnostic and prognostic modelling	72	See Supplementary file: TRIPOD guidelines
Comparison with known factors	73	Model introspection was performed to assess whether the tested features were responsible of the predictive performance
Multicollinearity	74	Hierarchical clustering of highly correlated features (absolute Spearman correlation coefficient ≥ 0.9). Features in each cluster are averages to create meta-features (2)
Model availability	75	Available on GitHub ¹
Data availability	76	Images not available. Extracted features available on GitHub ¹

¹ https://github.com/TimZaragori/Sklearn_NestedCV/tree/master/Radiomics_gliomas_article

² <https://github.com/Radiomics/pyradiomics>

1. Ginet M, Zaragori T, Marie P-Y, et al. Integration of dynamic parameters in the analysis of 18F-FDopa PET imaging improves the prediction of molecular features of gliomas. *Eur J Nucl Med Mol Imaging*. 2020;47:1381-1390.
2. Leger S, Zwanenburg A, Pilz K, et al. A comparative study of machine learning methods for time-to-event survival data for radiomics risk modelling. *Sci Rep*. 2017;7:1-11.

Supplementary Table 3. Set of hyperparameters tested for each algorithm. Tunable parameters are either a distribution with an interval from which 100 random values were selected or a list of tested values.

Algorithm	Default hyperparameters	Tunable hyperparameters
LR_L2		Regularization: log_uniform(10^{-3} ; 10^3) Size: int_uniform(2, 40)
Nnet	Number of hidden layers: 1	Regularization: log_uniform(10^{-4} ; 10^3) Learning rate: log_uniform(10^{-4} ; 10^{-2})
RF	Number of trees: 500	Max_features: 0.1, 0.2, ..., 0.8, 0.9 Min_samples_split: 0.1, 0.2, ..., 0.8, 0.9
SVM_RBF		Regularization: log_uniform(10^{-4} ; 10^4) Kernel coefficient: log_uniform(10^{-4} ; 10^4)
SVM_linear		Regularization: log_uniform(10^{-4} ; 10^4)

LR_L2: logistic regression with L2 regularization; Nnet: neural network; RF: random forest; SVM_RBF: support vector machine with a radial basis function kernel; SVM_linear: support vector machine with a linear kernel; log_uniform: logarithmic uniform continuous distribution; int_uniform: integer uniform discrete distribution; Max_features: maximum fraction of features to consider when looking for the best split; Min_samples_split: minimum fraction of samples required for splitting.

Supplementary Table 4. Adjusted p values of Mann-Whitney's tests for all features extracted from healthy striatum VOI for the conventional and the digital PET/CT before and after MComBat harmonization. Bold values are considered as significant.

Feature	Before MComBat	After MComBat
TBR _{mean}	1.33E-08	1
TBR _{max}	5.38E-08	1
TBR _{peak}	1.65E-07	1
TSR _{mean}	1	1
TSR _{max}	0.906553	1
TSR _{peak}	4.77E-05	1
MTV	7.38E-06	1
TTP	1	N.A.
Slope	0.790608814	N.A.
Volume (mesh-based) _{MORPH}	9.09727E-05	1
Volume (counting) _{MORPH}	7.06953E-06	1
Surface area _{MORPH}	0.000895079	1
Surface to volume ratio _{MORPH}	4.88916E-05	1
Sphericity _{MORPH}	0.002091582	1
Maximum 3D diameter _{MORPH}	0.081070391	1
Major axis length _{MORPH}	0.554057946	1
Minor axis length _{MORPH}	0.044208455	1
Least axis length _{MORPH}	1.38052E-06	1
Elongation _{MORPH}	0.206823797	1
Flatness _{MORPH}	1.8003E-06	1
Local intensity peak _{LI}	1.8003E-06	1
Global intensity peak _{LI}	1.29899E-07	1
Mean _{IS}	9.18584E-09	1
Variance _{IS}	0.103667981	1
Skewness _{IS}	0.905009215	1
Excess kurtosis _{IS}	0.353002802	1
Median _{IS}	6.74447E-09	1
Minimum _{IS}	0.00797326	1
10th percentile _{IS}	1.89334E-09	1
90th percentile _{IS}	8.38793E-08	1
Maximum _{IS}	1.32834E-08	1
Interquartile range _{IS}	0.067209156	1
Range _{IS}	4.39506E-06	1
Mean absolute deviation _{IS}	0.099686014	1
Robust mean absolute deviation _{IS}	0.095821679	1
Energy _{IS}	3.74004E-12	1
Root mean square _{IS}	1.01404E-08	1
Entropy _{IH}	0.151206065	1
Uniformity _{IH}	0.404042194	1
Joint maximum _{GLCM}	3.55154E-06	1
Joint average _{GLCM}	2.62298E-08	1
Joint variance _{GLCM}	0.381277285	1
Joint entropy _{GLCM}	8.68837E-07	1
Difference average _{GLCM}	3.74004E-12	1
Difference variance _{GLCM}	3.74004E-12	1
Difference entropy _{GLCM}	3.74004E-12	1

Sum average _{GLCM}	2.62298E-08	1
Sum entropy _{GLCM}	0.000570852	1
Angular second moment _{GLCM}	3.80425E-07	1
Contrast _{GLCM}	3.74004E-12	1
Dissimilarity _{GLCM}	3.74004E-12	1
Inverse difference _{GLCM}	3.74004E-12	1
Inverse difference normalised _{GLCM}	3.74004E-12	1
Inverse difference moment _{GLCM}	3.74004E-12	1
Inverse difference moment normalised _{GLCM}	3.74004E-12	1
Inverse variance _{GLCM}	1.34223E-09	1
Correlation _{GLCM}	3.74004E-12	1
Autocorrelation _{GLCM}	2.21647E-08	1
Cluster tendency _{GLCM}	4.88916E-05	1
Cluster shade _{GLCM}	0.000796999	1
Cluster prominence _{GLCM}	6.0706E-06	1
Information correlation 1 _{GLCM}	3.74004E-12	1
Information correlation 2 _{GLCM}	3.74004E-12	1
Short runs emphasis _{GLRLM}	3.74004E-12	1
Long runs emphasis _{GLRLM}	3.74004E-12	1
Low grey level run emphasis _{GLRLM}	1.06587E-08	1
High grey level run emphasis _{GLRLM}	1.39736E-08	1
Short run low grey level emphasis _{GLRLM}	1.31002E-11	1
Short run high grey level emphasis _{GLRLM}	0.016928903	1
Long run low grey level emphasis _{GLRLM}	1.28729E-08	1
Long run high grey level emphasis _{GLRLM}	3.74004E-12	1
Grey level non-uniformity _{GLRLM}	3.74004E-12	1
Grey level non-uniformity normalised _{GLRLM}	0.95622021	1
Run length non-uniformity _{GLRLM}	3.74004E-12	1
Run length non-uniformity normalised _{GLRLM}	3.74004E-12	1
Run percentage _{GLRLM}	3.74004E-12	1
Grey level variance _{GLRLM}	0.337884484	1
Run length variance _{GLRLM}	3.74004E-12	1
Run entropy _{GLRLM}	7.55873E-12	1
Small zone emphasis _{GLSZM}	7.5676E-05	1
Large zone emphasis _{GLSZM}	3.74004E-12	1
Low grey level emphasis _{GLSZM}	4.79732E-07	1
High grey level emphasis _{GLSZM}	3.80425E-07	1
Small zone low grey level emphasis _{GLSZM}	3.46597E-07	1
Small zone high grey level emphasis _{GLSZM}	0.00409295	1
Large zone low grey level emphasis _{GLSZM}	3.74004E-12	1
Large zone high grey level emphasis _{GLSZM}	3.74004E-12	1
Grey level non-uniformity _{GLSZM}	0.917448067	1
Grey level non-uniformity normalised _{GLSZM}	0.072949459	1
Zone size non-uniformity _{GLSZM}	0.00797326	1
Zone size non-uniformity normalised _{GLSZM}	3.20696E-05	1
Zone percentage _{GLSZM}	3.74004E-12	1
Grey level variance _{GLSZM}	0.003426848	1
Zone size variance _{GLSZM}	3.74004E-12	1
Zone size entropy _{GLSZM}	0.067209156	1
Coarseness _{NGTDM}	3.74004E-12	1
Contrast _{NGTDM}	3.74004E-12	1

Busyness _{NGTDM}	4.08881E-12	1
Complexity _{NGTDM}	0.000448647	1
Strength _{NGTDM}	5.12053E-12	1
Low dependence emphasis _{NGLDM}	3.74004E-12	1
High dependence emphasis _{NGLDM}	3.74004E-12	1
Low grey level count emphasis _{NGLDM}	7.60045E-09	1
High grey level count emphasis _{NGLDM}	1.01404E-08	1
Low dependence low grey level emphasis _{NGLDM}	3.74004E-12	1
Low dependence high grey level emphasis _{NGLDM}	3.85772E-12	1
High dependence low grey level emphasis _{NGLDM}	3.85772E-12	1
High dependence high grey level emphasis _{NGLDM}	3.74004E-12	1
Grey level non-uniformity _{NGLDM}	3.74004E-12	1
Dependence count non-uniformity _{NGLDM}	3.74004E-12	1
Dependence count non-uniformity normalised _{NGLDM}	3.74004E-12	1
Grey level variance _{NGLDM}	0.115778585	1
Dependence count variance _{NGLDM}	3.74004E-12	1

Supplementary Table 5. Adjusted p values of Mann-Whitney's tests for all normalized features extracted from tumor VOI for carbidopa premedicated (n=42) and non-premedicated (n=30) patients.

Feature	With Carbidopa vs Without Carbidopa
TBR _{mean}	0,995171012
TBR _{max}	0,995171012
TBR _{peak}	0,995171012
TSR _{mean}	0,995171012
TSR _{max}	0,995171012
TSR _{peak}	0,995171012
MTV	0,995171012
TTP	0,913139725
Slope	0,992539943
Volume (mesh-based) _{MORPH}	0,995171012
Volume (counting) _{MORPH}	0,995171012
Surface area _{MORPH}	0,995171012
Surface to volume ratio _{MORPH}	0,995171012
Sphericity _{MORPH}	0,995171012
Maximum 3D diameter _{MORPH}	0,995171012
Major axis length _{MORPH}	0,995171012
Minor axis length _{MORPH}	0,995171012
Least axis length _{MORPH}	0,995171012
Elongation _{MORPH}	0,995171012
Flatness _{MORPH}	0,995171012
Local intensity peak _{LI}	0,995171012
Global intensity peak _{LI}	0,995171012
Mean _{IS}	0,995171012
Variance _{IS}	0,995171012
Skewness _{IS}	0,995171012
Excess kurtosis _{IS}	0,995171012
Median _{IS}	0,995171012
Minimum _{IS}	0,995171012
10th percentile _{IS}	0,995171012
90th percentile _{IS}	0,995171012
Maximum _{IS}	0,995171012
Interquartile range _{IS}	0,995171012
Range _{IS}	0,995171012
Mean absolute deviation _{IS}	0,995171012
Robust mean absolute deviation _{IS}	0,995171012
Energy _{IS}	0,995171012
Root mean square _{IS}	0,995171012
Entropy _{IH}	0,995171012
Uniformity _{IH}	0,995171012
Joint maximum _{GLCM}	0,995171012
Joint average _{GLCM}	0,995171012
Joint variance _{GLCM}	0,995171012
Joint entropy _{GLCM}	0,995171012
Difference average _{GLCM}	0,995171012

Difference variance _{GLCM}	0,995171012
Difference entropy _{GLCM}	0,995171012
Sum average _{GLCM}	0,995171012
Sum entropy _{GLCM}	0,995171012
Angular second moment _{GLCM}	0,995171012
Contrast _{GLCM}	0,995171012
Dissimilarity _{GLCM}	0,995171012
Inverse difference _{GLCM}	0,995171012
Inverse difference normalised _{GLCM}	0,995171012
Inverse difference moment _{GLCM}	0,995171012
Inverse difference moment normalised _{GLCM}	0,995171012
Inverse variance _{GLCM}	0,995171012
Correlation _{GLCM}	0,995171012
Autocorrelation _{GLCM}	0,995171012
Cluster tendency _{GLCM}	0,995171012
Cluster shade _{GLCM}	0,995171012
Cluster prominence _{GLCM}	0,995171012
Information correlation 1 _{GLCM}	0,995171012
Information correlation 2 _{GLCM}	0,995171012
Short runs emphasis _{GLRLM}	0,995171012
Long runs emphasis _{GLRLM}	0,995171012
Low grey level run emphasis _{GLRLM}	0,995171012
High grey level run emphasis _{GLRLM}	0,995171012
Short run low grey level emphasis _{GLRLM}	0,995171012
Short run high grey level emphasis _{GLRLM}	0,995171012
Long run low grey level emphasis _{GLRLM}	0,995171012
Long run high grey level emphasis _{GLRLM}	0,995171012
Grey level non-uniformity _{GLRLM}	0,995171012
Grey level non-uniformity normalised _{GLRLM}	0,995171012
Run length non-uniformity _{GLRLM}	0,995171012
Run length non-uniformity normalised _{GLRLM}	0,995171012
Run percentage _{GLRLM}	0,995171012
Grey level variance _{GLRLM}	0,995171012
Run length variance _{GLRLM}	0,995171012
Run entropy _{GLRLM}	0,995171012
Small zone emphasis _{GLSZM}	0,995171012
Large zone emphasis _{GLSZM}	0,995171012
Low grey level emphasis _{GLSZM}	0,995171012
High grey level emphasis _{GLSZM}	0,995171012
Small zone low grey level emphasis _{GLSZM}	0,995171012
Small zone high grey level emphasis _{GLSZM}	0,995171012
Large zone low grey level emphasis _{GLSZM}	0,995171012
Large zone high grey level emphasis _{GLSZM}	0,995171012
Grey level non-uniformity _{GLSZM}	0,995171012
Grey level non-uniformity normalised _{GLSZM}	0,995171012
Zone size non-uniformity _{GLSZM}	0,995171012
Zone size non-uniformity normalised _{GLSZM}	0,995171012
Zone percentage _{GLSZM}	0,995171012
Grey level variance _{GLSZM}	0,995171012
Zone size variance _{GLSZM}	0,995171012
Zone size entropy _{GLSZM}	0,995171012

Coarseness _{NGTDM}	0,995171012
Contrast _{NGTDM}	0,995171012
Busyness _{NGTDM}	0,995171012
Complexity _{NGTDM}	0,995171012
Strength _{NGTDM}	0,995171012
Low dependence emphasis _{NGLDM}	0,995171012
High dependence emphasis _{NGLDM}	0,995171012
Low grey level count emphasis _{NGLDM}	0,995171012
High grey level count emphasis _{NGLDM}	0,995171012
Low dependence low grey level emphasis _{NGLDM}	0,995171012
Low dependence high grey level emphasis _{NGLDM}	0,995171012
High dependence low grey level emphasis _{NGLDM}	0,995171012
High dependence high grey level emphasis _{NGLDM}	0,995171012
Grey level non-uniformity _{NGLDM}	0,995171012
Dependence count non-uniformity _{NGLDM}	0,995171012
Dependence count non-uniformity normalised _{NGLDM}	0,995171012
Grey level variance _{NGLDM}	0,995171012
Dependence count variance _{NGLDM}	0,995171012

Supplementary Figure 1. Heatmap depicting the mean area under the curve of the receiver operator characteristic in outer loops on validation sets as a function of the number of features selected (columns) and classification algorithm (rows) for the IDH mutations classification task.

LR_L2: logistic regression with L2 regularization; Nnet: neural network; RF: random forest; SVM_RBF: support vector machine with a radial basis function kernel; SVM_linear: support vector machine with a linear kernel; wlcx_score_5: 5 features selected using Wilcoxon score; wlcx_score_10: 10 features selected using Wilcoxon score; wlcx_score_15: 15 features selected using Wilcoxon score.

Supplementary Figure 2. Heatmap depicting the mean area under the curve of the receiver operator characteristic in outer loops on validation sets as a function of the number of features selected (columns) and classification algorithm (rows) for the 1p/19q co-deletion classification task.

LR_L2: logistic regression with L2 regularization; Nnet: neural network; RF: random forest; SVM_RBF: support vector machine with a radial basis function kernel; SVM_linear: support vector machine with a linear kernel; wlcx_score_5: 5 features selected using Wilcoxon score; wlcx_score_10: 10 features selected using Wilcoxon score; wlcx_score_15: 15 features selected using Wilcoxon score.

Supplementary Figure 3. Results of performance comparisons between datasets including only static parameters (S), static and dynamic parameters (S+D), static and other radiomics features (S+OR) and the 114-feature dataset (S+D+OR). The figures present mean (full line) and 95% confidence intervals (dotted line) of the 50-outer area under the receiver operating characteristic curves (AUC) with the best model of each dataset for the classification of A. the IDH mutations and B. the 1p/19q co-deletion. Kruskal Wallis tests and paired comparisons were performed to compare the 4 datasets to each other. ¥ denotes significantly better performances over S (mean outer AUC increase of 39% with S+D+OR, 37% with S+D and 22% with S+OR for the prediction of the IDH mutations, 21% with S+D+OR and 22% with S+OR for the prediction of the 1p/19q co-deletion). † denotes significantly better performances over S+OR (mean outer AUC increase of 14% with S+D+OR and 12% with S+D for the prediction of the IDH mutations). § denotes significantly better performances over S+D (mean outer AUC increase of 37% with S+D+OR and 38% with S+OR for the prediction of the 1p/19q co-deletion). The best model used in A. was the logistic regression and 5 features selected. In B. the best model used was the support vector machine with a radial basis function kernel and 15 features selected.