




## Article

# Outcome Prediction at Patient Level Derived from Pre-Treatment 18F-FDG PET Due to Machine Learning in Metastatic Melanoma Treated with Anti-PD1 Treatment

Anthime Flaus<sup>1,2,\*</sup>, Vincent Habouzit<sup>1</sup>, Nicolas de Leiris<sup>3,4</sup> , Jean-Philippe Vuillez<sup>3,4</sup>, Marie-Thérèse Leccia<sup>5</sup>, Mathilde Simonson<sup>6</sup>, Jean-Luc Perrot<sup>7</sup> , Florent Cachin<sup>6</sup> and Nathalie Prevot<sup>1,8</sup> 

- <sup>1</sup> Nuclear Medicine Department, Saint-Etienne University Hospital, University of Saint-Etienne, 42000 Saint-Etienne, France; vincent.habouzit@gmail.com (V.H.); nathalie.prevot@chu-st-etienne.fr (N.P.)
- <sup>2</sup> Nuclear Medicine Department, Hospices Civils de Lyon, University of Lyon, 69008 Lyon, France
- <sup>3</sup> Nuclear Medicine Department, CHU Grenoble Alpes, University Grenoble Alpes, 38000 Grenoble, France; ndeleiris@chu-grenoble.fr (N.d.L.); jpvuillez@chu-grenoble.fr (J.-P.V.)
- <sup>4</sup> INSERM, CHU Grenoble Alpes, Laboratoire Radiopharmaceutiques Biocliniques, University Grenoble Alpes, 38000 Grenoble, France
- <sup>5</sup> Dermatology Department, CHU Grenoble Alpes, University Grenoble Alpes, 38000 Grenoble, France; mtleccia@chu-grenoble.fr
- <sup>6</sup> Jean Perrin Cancer Centre of Clermont-Ferrand, Nuclear Medicine Department, 63011 Clermont-Ferrand, France; mathildes@startmail.com (M.S.); florent.cachin@clermont.unicancer.fr (F.C.)
- <sup>7</sup> Dermatology Department, Saint-Etienne University Hospital, University of Saint-Etienne, 42000 Saint-Etienne, France; j.luc.perrot@chu-st-etienne.fr
- <sup>8</sup> INSERM U 1059 Sainbiose, University of Saint-Etienne, 42000 Saint-Etienne, France
- \* Correspondence: anthime.flaus@gmail.com; Tel.: +33-477-828-318



**Citation:** Flaus, A.; Habouzit, V.; de Leiris, N.; Vuillez, J.-P.; Leccia, M.-T.; Simonson, M.; Perrot, J.-L.; Cachin, F.; Prevot, N. Outcome Prediction at Patient Level Derived from Pre-Treatment 18F-FDG PET Due to Machine Learning in Metastatic Melanoma Treated with Anti-PD1 Treatment. *Diagnostics* **2022**, *12*, 388. <https://doi.org/10.3390/diagnostics12020388>

Academic Editor: Barbara Palumbo

Received: 17 December 2021

Accepted: 30 January 2022

Published: 2 February 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** (1) Background: As outcome of patients with metastatic melanoma treated with anti-PD1 immunotherapy can vary in success, predictors are needed. We aimed to predict at the patients' levels, overall survival (OS) and progression-free survival (PFS) after one year of immunotherapy, based on their pre-treatment 18F-FDG PET; (2) Methods: Fifty-six metastatic melanoma patients—without prior systemic treatment—were retrospectively included. Forty-five 18F-FDG PET-based radiomic features were computed and the top five features associated with the patient's outcome were selected. The analyzed machine learning classifiers were random forest (RF), neural network, naive Bayes, logistic regression and support vector machine. The receiver operating characteristic curve was used to compare model performances, which were validated by cross-validation; (3) Results: The RF model obtained the best performance after validation to predict OS and PFS and presented AUC, sensitivities and specificities (IC95%) of  $0.87 \pm 0.1$ ,  $0.79 \pm 0.11$  and  $0.95 \pm 0.06$  for OS and  $0.9 \pm 0.07$ ,  $0.88 \pm 0.09$  and  $0.91 \pm 0.08$  for PFS, respectively. (4) Conclusion: A RF classifier, based on pretreatment 18F-FDG PET radiomic features may be useful for predicting the survival status for melanoma patients, after one year of a first line systemic treatment by immunotherapy.

**Keywords:** 18F-FDG PET; radiomics; machine-learning; immunotherapy; melanoma; prediction

## 1. Introduction

Immunotherapy using immune checkpoint inhibitors (ICI) that target programmed cell death 1 receptor (anti-PD1) has substantially changed the therapeutic strategies for cancers such as metastatic melanoma [1]. However, as only subsets of patients will take advantage from it, it is necessary to develop non-invasive tools to stratify pre-treatment risk and assess prognosis, in order to prevent toxicities and hasten the introduction of more appropriate treatments. Nevertheless, no tools are currently clinically identified as helpful to select patients and which tools to use remains an open question [2]. Moreover, traditional cohort-oriented methods (such as the Kaplan–Meier survival techniques [3] to investigate real-world evidence data) have shown limited results because of their difficulties

to cope with heterogeneous patient populations and their inability to provide patient-level predictions [4,5].

Computational medical imaging—known as radiomics—involves the analysis and translation of medical images into quantitative data. It allows an in-depth characterization of tumor imaging phenotypes with the underlying hypothesis that images reflect not only the tissues' macroscopic but also their cellular and molecular properties. These image driven biomarkers could serve as tools to better aid clinical decisions [6].

Glucose analog, 2-deoxy-2-[<sup>18</sup>F]fluoro-D-glucose (18F-FDG) uptake represents glucose metabolic activity and is used as a tracer of positron emission tomography (PET)/X-ray computed tomography (CT) for initial staging in cases of high-risk melanoma, macroscopic lymph node involvement and/or known distant metastases [7].

Few studies have examined the prognostic value of 18F-FDG PET radiomic features for metastatic melanoma before immunotherapy treatment, but associations have been found between 18F-FDG PET parameters and outcome at a group level [8–10]. However, to our knowledge, no study has previously investigated patient-level prediction performances before treating them with anti-PD1 as a first line therapy. Therefore, our aim was to assess the prediction performance of a pre-treatment 18F-FDG PET-based machine learning model for separating individual patients with metastatic melanoma, according to overall survival (OS) and progression-free survival (PFS), after one year of a first line systemic treatment, by immunotherapy (anti-PD1)

## 2. Materials and Methods

The protocol was approved (2 July 2020) by the institutional medical ethics committee (IRB: IORG0007394) of Saint-Etienne (IRBN 842020/CHUSTE) and all methods were carried out in accordance with relevant guidelines and regulations. Informed consent to participate in the study was obtained from all the patients according to national regulations.

### 2.1. Patients

The hospital information system of two university hospitals (Saint-Etienne and Grenoble) was investigated to identify metastatic melanoma patients treated by a first line of anti-PD1 treatment and imaged with an 18F-FDG PET-CT scan before therapy, between January 2016 and January 2019. Electronic clinical and radiological databases were used to obtain patients' demographical details, clinical history and metastatic status—as defined by the eighth edition American Joint Committee on Cancer (AJCC) melanoma staging system [11], anatomopathological data, treatment data, clinical outcome and follow-up duration as well as 18F-FDG PET data.

Out of 64 screened patients, 56 were included. Inclusion criteria were as follows: (1) metastatic melanoma proven by biopsy; (2) anti-PD1 antibodies were the first line of treatment, with no previous systematic therapy; (3) patients were B-RAF wild type; and (4) pre-treatment 18F-FDG PET data were available. Eight patients were excluded: (1) patients with no measurable disease or no significant FDG avid tumor ( $n = 6$ ), including patients presenting only brain metastases due to the impossibility of clearly delineating their metabolic tumoral volume; and (2) patients with lesion sizes inferior to 64 voxels ( $n = 2$ ), since the radiomic analysis becomes irrelevant below this threshold [12,13].

### 2.2. Follow-Up of Patients

Overall survival, progression-free survival or time to last censoring were recorded starting from the date of the PET scan. OS was defined as the time (in months) between the PET scan and the date of death, due to any cause. PFS was defined as the time between the PET scan and locoregional or distant relapse, or death from any cause. Living patients were censored at the time of the last clinical follow-up. Patient response was assessed with Response Evaluation Criteria in Solid Tumors (RECIST 1.1) using diagnostic contrast-enhanced CT and all available clinical information.

### 2.3. Imaging Protocols

Scans were acquired using different PET-CT systems: Biograph mCT Flow 20, Biograph 6 HI-REZ and Biograph Horizon 16 (Siemens Healthcare, Erlangen, Germany) and Discovery 690 (General Electric Healthcare, Chicago, IL, USA). Patients fasted for at least 6 h and blood glucose levels were confirmed to be under 180 mg/dl. Patients were injected according to current guidelines with an activity of 2.5–4 MBq/kg  $^{18}\text{F}$ -FDG (median activity 269 MBq, range 146–468 MBq). Sixty minutes after injection, whole-body PET and unenhanced CT images were acquired. The CT attenuation-corrected acquired data were reconstructed using a three-dimensional ordered subset expectation–maximization algorithm as previously described [10]. Each scanner used consistent reconstruction settings, matrix and voxel size. Each voxel in PET images were converted into standard uptake value (SUV) [14].

### 2.4. Radiomics: Feature Extraction

A feature extraction process was performed using the free software LIFEx (v4.0, Local Image Feature Extraction) which is compliant with the image biomarker standardization initiative (IBSI) [13,15]. One experienced nuclear medicine physician analyzed and segmented  $^{18}\text{F}$ -FDG PET scans without knowledge of the patients' clinical outcome. A semi-automatic method was employed for segmentation based on a 40% SUVmax threshold [16]. Each hypermetabolic lesion was segmented to create a volume of interest (VOI) that was equivalent to the metabolic tumoral volume (MTV).

For each patient, radiomic features were extracted from the VOI of the lesion with the highest  $^{18}\text{F}$ -FDG uptake. This was consistent with the Positron Emission Tomography Response Criteria in Solid Tumors (PERCIST) definition of target lesion [17]. A total of 45 features were computed for each VOI after voxel intensity resampling using 64 discrete values between 0 and 32 SUV units [18].

The 14 first order features were: 4 features from shape, 5 from histogram and 5 from conventional indexes (SUVmin, SUVmean, SUVmax, SUVpeak and total lesion glycolysis or TLG). The 31 second order textural features (derived from different matrices) were: 6 from the grey level co-occurrence matrix (GLCM), 11 from the grey-level run length matrix (GLRLM), 3 from the neighborhood grey-level different matrix (NGLDM) and 11 from the grey-level zone length matrix (GLZLM). A detailed description of each parameter is available in the technical appendix of the LIFEx software [13].

### 2.5. Radiomics: Feature Harmonization

Since different PET scanners were used, post-reconstruction harmonization method was performed for all PET parameters using the ComBat harmonization method in R with a software package (<https://github.com/Jfortin1/ComBatHarmonization> (accessed on 5 August 2020)). ComBat harmonization was described in 2007 [19] and is now widely used in genomics: its efficacy in PET has been shown [20].

### 2.6. Radiomics: Feature Selection

A feature reduction procedure is necessary to select a subset of useful features that increase the prediction accuracy and reduce overfitting [21]. First, multicollinearity of the features were analyzed by Spearman correlation, with a threshold for the correlation coefficient of 0.8 [22]. Secondly, information gain ratio [23], Gini index [21] and F-score [24] were used as ranking and filter-based feature selection methods to reduce feature dimension. Radiomic features were ranked based on the three selection methods according to their correlation with survival status at 1-year follow-up. A total of 5 best-ranked features were selected based on the overall ranks across the 3 scoring methods [25]. Selected features were standardized using the standard scaler function to unit variance within the training set and the same scaler was used to transform the test set. Synthetic minority oversampling technique was used to correct for imbalanced datasets [26].

### 2.7. Machine Learning Approach for Classification

Five different machine learning algorithms used as classification models were compared, including neural network, logistic regression, support vector machine, random forest and naïve bayes. Given that the difference between our defined binary response and predicted response by each classification model can be described by a confusion matrix, we can define the following properties: the number of true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN). The models' performances were then evaluated by calculating the area under the curve (AUC) of the receiver operating characteristic (ROC) curve and mapping the models' sensitivity and specificity measured by the TPR and 1-TNR, respectively [27]. Each model's performance was validated using a sampling technique method: "cross-validation". The patients' data were randomly separated into training sets and testing sets stratified on the outcome (training set size: 75%). The whole procedure was repeated 50 times and results were averaged.

The statistical analyses were performed using the open-source *R* software package [28] (version 3.0.1, <http://www.Rproject.org> (accessed on 1 August 2020)). A Mann–Whitney U test or chi-square test enabled us to appropriately assess the difference between two quantitative variables or compare categorical data. The machine learning approach was performed using Python (version 3.7) and the open source Numpy, Scipy and Scikit-learn packages. All *p* values are two-sided, and a *p* value < 0.05 is considered statistically significant.

## 3. Results

### 3.1. Patient Characteristics

A total of 56 patients were included with a median age of 68 years old (range: 40–84). The median follow-up was 22.1 months (range: 2.1–49.2). At one year, 28 patients (50%) presented progressive diseases and 16 patients (29%) had died. We found no statistically significant differences for any parameters—sex, age, university hospital, metastatic status, before starting anti-PD1 treatment, histological characteristics of the initial melanoma, localization or initial cancer staging—between patients alive or dead at one year, patients with progressive disease versus partial or complete response or patients with stable disease.

### 3.2. Selection of Radiomic Features for Prediction

We ranked selected radiomic features according to the scoring methods, as shown in Table 1. For OS, the top five predictive parameters were: GLZLM long-zone emphasis, GLZLM long-zone high grey-level emphasis, GLZLM zone length non-uniformity, GLCM homogeneity and histogram kurtosis. For PFS, they were: GLZLM long-zone emphasis, GLZLM long-zone low grey-level emphasis, GLZLM long-zone high grey-level emphasis CONVENTIONAL SUV standard deviation and GLCM entropy log10. Parameters extracted from GLZLM were the most informative to predict OS and PFS in our cohort.

**Table 1.** Ranking of the selected radiomic features extracted from the 18F-FDG PET, according to association with survival at one year.

Overall Survival		Progression Free Survival	
Overall Rank	Radiomic Feature	Overall Rank	Radiomic Feature
1	GLZLM_LZE	1	GLZLM_LZE
2	GLZLM_LZHGE	2	GLZLM_LZLGE
3	GLZLM_ZLNU	3	GLZLM_LZHGE
4	GLCM_Homogeneity	4	CONVENTIONAL_SUVstd
5	HISTO_Kurtosis	5	GLCM_Entropy_log10

Table 1. Cont.

Overall Survival		Progression Free Survival	
Overall Rank	Radiomic Feature	Overall Rank	Radiomic Feature
6	GLRLM_GLNU	6	NGLDM_Coarseness
7	GLRLM_LRHGE	7	HISTO_Skewness
8	NGLDM_Coarseness	8	CONVENTIONAL_TLG
9	HISTO_Energy	9	HISTO_Entropy_log10
10	GLRLM_SRE	10	GLCM_Energy
11	HISTO_Entropy_log10	11	CONVENTIONAL_SUVmax
12	NGLDM_Contrast	12	GLRLM_GLNU
13	GLCM_Correlation	13	GLRLM_LRHGE
14	GLRLM_LGRE		
15	GLZLM_SZE		
16	HISTO_Skewness		
17	GLCM_Contrast		
18	CONVENTIONAL_SUVmin		

NGLDM: neighborhood grey-level different matrix, GLZLM: grey-level zone length matrix, GLRLM: grey-level run length matrix, HISTO: histogram, LZHGE: long-zone high grey-level emphasis; LZE: long-zone emphasis; LZLGE: long-zone low grey-level emphasis RLNU: run length non-uniformity, ZLNU: zone length non-uniformity; TLG: total lesion glycolysis, GLNU: grey-level non-uniformity, LRHGE: long-run high grey-level emphasis, GLCM: grey-level co-occurrence matrix, SZE: short-zone emphasis, SUV: standard uptake value, std: standard deviation.

### 3.3. Performances of the Five Machine Learning Methods for Patient Classification

The overall classification performances of the five models for OS and PFS in the validation set are shown in Table 2. The random forest model obtained the best performance for both OS and PFS. AUC (95% CI), sensitivities (95% CI) and specificities (95% CI) were  $0.87 \pm 0.1$ ,  $0.79 \pm 0.11$  and  $0.95 \pm 0.06$  for OS and  $0.9 \pm 0.07$ ,  $0.88 \pm 0.09$  and  $0.91 \pm 0.08$  for PFS, respectively.

**Table 2.** Performance of each model using cross-validation to predict overall survival and progression-free survival at one year, in patients with metastatic melanoma.

	Overall survival				
	NB	LR	RF	SVM	NN
AUC (95% CI)	$0.82 \pm 0.15$	$0.84 \pm 0.15$	$0.87 \pm 0.1$	$0.82 \pm 0.15$	$0.84 \pm 0.14$
Sensitivity (95% CI)	$0.78 \pm 0.11$	$0.81 \pm 0.11$	$0.79 \pm 0.11$	$0.77 \pm 0.12$	$0.79 \pm 0.11$
Specificity (95% CI)	$0.86 \pm 0.1$	$0.87 \pm 0.10$	$0.95 \pm 0.06$	$0.87 \pm 0.09$	$0.89 \pm 0.09$
	Progression-free survival				
	NB	LR	RF	SVM	NN
AUC (95% CI)	$0.69 \pm 0.15$	$0.64 \pm 0.14$	$0.90 \pm 0.07$	$0.63 \pm 0.15$	$0.69 \pm 0.13$
Sensitivity (95% CI)	$0.80 \pm 0.11$	$0.59 \pm 0.14$	$0.88 \pm 0.09$	$0.55 \pm 0.14$	$0.58 \pm 0.14$
Specificity (95% CI)	$0.57 \pm 0.14$	$0.70 \pm 0.12$	$0.91 \pm 0.08$	$0.72 \pm 0.12$	$0.81 \pm 0.11$

RF: random forest, LR: logistic regression, NN: neural network, SVM: support vector classification, NB: naïve bayes.

## 4. Discussion

This study intended to predict survival status using pre-therapeutic  $^{18}\text{F}$ -FDG PET of metastatic melanoma patients, after one year of a first line systemic treatment by immunotherapy (anti-PD1). Five radiomic features, extracted from  $^{18}\text{F}$ -FDG PET, which captured information about tumor phenotype, were used to develop and validate five machine learning models. Using a cross validation method, we found the random forest classifier model predicted PFS and OS at one year with the highest accuracy.



In this study, the five selected radiomics features were second order features, which possessed stronger correlations with outcome in this cohort than conventional 18F-FDG-PET metrics. Indeed, conventional metrics, i.e., SUVmax, SUVmean and MTV, were initially included in the extracted radiomics set but were not selected for analysis due to their lower contributions to the outcome prediction. The choice of the best radiomics parameters from the pretherapeutic 18F-FDG-PET to predict the outcome is still a matter of debate as results from previous studies are heterogeneous. One study reporting on SUVmax found SUVmean and tumor heterogeneity index were not correlated with OS or PFS in a heterogeneous cohort of 55 melanoma patients before anti-PD1 treatment [9]; however, other studies showed significant correlations between conventional or second order 18F-FDG PET parameters and melanoma patients' outcomes after immunotherapy [8,10,29]. Authors showed conventional parameters, such as SUV max, SUVpeak, TLG and MTV, were associated with OS [8,10,29]. Second order features, such as tumor heterogeneity index, as well as long-zone emphasis from the GLZLM matrix were associated with OS [8,10]. Moreover, six second order features correlated with treatment response, among whom three came from the GLZLM matrix [29]. This is coherent with our ranking of radiomic features to predict PFS, which is best with three GLZLM features.

In line with previous studies, the model employing the random forest classifier—known as a classifier with high ability to learn and predict in small cohorts [30]—obtained the best performance among the machine learning classifiers in our cohort. In other studies, this classifier, based on radiomics, predicted disease recurrence accurately for non-small cell lung cancer [25], enabled risk stratification of gastro-intestinal stromal tumors [31] and helped differentiate recurrent brain tumors from radiation necrosis [32].

Previous studies have shown at group levels that 18F-FDG-PET may help discriminate melanoma patients that will benefit from immunotherapy versus patients with poor prognosis. The most informative parameters are total metabolic tumoral volume, bone to liver uptake ratio and two radiomics features: tumor heterogeneity index and GLZLM long-zone emphasis [8–10]. However, even if this information is of great importance, we can still refine our patient-level risk stratification. Our results using a classifier provide an additional biomarker. Similar results are available in non-small cell lung cancer [33].

Nevertheless, our study had some limitations. Firstly, its retrospective design and the descriptive nature of a small number of cases may limit its generalization to other cases. Secondly, as group level analyses suggested, clinical parameters were not significantly associated with outcome, so we only used pure radiomic modeling and did not integrate clinical parameters for further analysis and comparison [9]. Thirdly, we could not consider immune RECIST (iRECIST) as an endpoint for PFS since it would have introduced a chronological bias. Indeed, the RECIST working group developed iRECIST criteria in 2017 [34], while our data were gathered starting from January 2016. Recent studies have however demonstrated a relative agreement between RECIST and iRECIST [35] and no case of pseudo-progression was reported in our cohort. Fourthly, the use of different scanners may have affected the results of the 18F-FDG PET-based radiomic analyses [36]. However, in our work, intensity binning was performed on all data to increase the adherence to key methodological principles of radiomics analysis. Moreover, we conducted post-reconstruction harmonization using the ComBat method [20] to correct the effect of multi PET scanner as performed in other recent studies; this also underlined the feasibility of multicentric radiomics studies, which are scarce [37,38]. Finally, we did not examine the pathological correlations of these radiomic features, which are largely unknown and remain to be clarified in the future.

## 5. Conclusions

To conclude, a random forest classifier based on 18F-FDG PET radiomic features may be useful to predict the survival status of metastatic melanoma patients after one year of treatment with immunotherapy (anti-PD1) as the first line choice. Such a biomarker presents the potential to help predict outcomes at the patient level and to improve risk stratification and treatment planning. A multicentric, prospective study, including more

patients, should be performed to confirm the predictive performance of this biomarker. Following this first validation, a clinical trial with personalized treatment according to the 18F-FDG PET biomarker should be investigated.

**Author Contributions:** A.F.: data curation, formal analysis conceptualization, methodology, validation and writing—original draft preparation; V.H. and N.d.L.: data curation, writing review and editing; N.P.: supervision, review and editing; J.-P.V. and M.S.: review and editing; F.C.: methodology, supervision, validation, review and editing; J.-L.P. and M.-T.L.: clinical evaluations, treatments, follow-up evaluations of the patient. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no funding.

**Institutional Review Board Statement:** The protocol was approved (2 July 2020) by the institutional medical ethics committee (IRB: IORG0007394) of Saint-Etienne (IRBN 842020/CHUSTE) and all methods were carried out in accordance with relevant guidelines and regulations. Informed consent to participate in the study was obtained from all the patients according to national regulations.

**Informed Consent Statement:** Informed consent to participate in the study was obtained from all the patients according to national regulations.

**Data Availability Statement:** The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request. The data are not publicly available due to ethical restriction.

**Acknowledgments:** The authors would like to thank Orhac for constructive criticism of the manuscript. They also want to express their gratitude to all the members of the PET staff for their contribution in performing this study.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Robert, C.; Long, G.V.; Brady, B.; Dutriaux, C.; Maio, M.; Mortier, L.; Hassel, J.C.; Rutkowski, P.; McNeil, C.; Kalinka-Warzocho, E.; et al. Nivolumab in Previously Untreated Melanoma without *BRAF* Mutation. *N. Engl. J. Med.* **2015**, *372*, 320–330. [[CrossRef](#)] [[PubMed](#)]
2. Buder-Bakhaya, K.; Hassel, J.C. Biomarkers for Clinical Benefit of Immune Checkpoint Inhibitor Treatment—a Review from the Melanoma Perspective and Beyond. *Front. Immunol.* **2018**, *9*, 1474. [[CrossRef](#)] [[PubMed](#)]
3. *Textbook of Pharmacoepidemiology*; Strom, B.L.; Kimmel, S.E. (Eds.) John Wiley & Sons: Chichester, UK/Hoboken, NJ, USA, 2006; ISBN 978-1-119-41341-7.
4. Ankarfeldt, M.; Adalsteinsson, E.; Groenwold, R.; Ali, M.S.; Klungel, O. A Systematic Literature Review on the Efficacy-Effectiveness Gap: Comparison of Randomized Controlled Trials and Observational Studies of Glucose-Lowering Drugs. *CLEP* **2017**, *9*, 41–51. [[CrossRef](#)] [[PubMed](#)]
5. Arora, P.; Boyne, D.; Slater, J.J.; Gupta, A.; Brenner, D.R.; Druzdzal, M.J. Bayesian Networks for Risk Prediction Using Real-World Data: A Tool for Precision Medicine. *Value Health* **2019**, *22*, 439–445. [[CrossRef](#)]
6. Gillies, R.J.; Kinahan, P.E.; Hricak, H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology* **2016**, *278*, 563–577. [[CrossRef](#)]
7. Salaün, P.-Y.; Abgral, R.; Malard, O.; Querellou-Lefranc, S.; Quere, G.; Wartski, M.; Coriat, R.; Hindie, E.; Taieb, D.; Tabarin, A.; et al. Good Clinical Practice Recommendations for the Use of PET/CT in Oncology. *Eur. J. Nucl. Med. Mol. Imaging* **2020**, *47*, 28–50. [[CrossRef](#)]
8. Sanli, Y.; Leake, J.; Odu, A.; Xi, Y.; Subramaniam, R.M. Tumor Heterogeneity on FDG PET/CT and Immunotherapy: An Imaging Biomarker for Predicting Treatment Response in Patients with Metastatic Melanoma. *AJR Am. J. Roentgenol.* **2019**, *212*, 1318–1326. [[CrossRef](#)]
9. Seban, R.-D.; Nemer, J.S.; Marabelle, A.; Yeh, R.; Deutsch, E.; Ammari, S.; Moya-Plana, A.; Mokrane, F.-Z.; Gartrell, R.D.; Finkel, G.; et al. Prognostic and Theranostic 18F-FDG PET Biomarkers for Anti-PD1 Immunotherapy in Metastatic Melanoma: Association with Outcome and Transcriptomics. *Eur. J. Nucl. Med. Mol. Imaging* **2019**, *46*, 2298–2310. [[CrossRef](#)]
10. Flaus, A.; Habouzit, V.; De Leiris, N.; Vuillez, J.P.; Leccia, M.T.; Perrot, J.L.; Prevot, N.; Cachin, F. FDG PET Biomarkers for Prediction of Survival in Metastatic Melanoma Prior to Anti-PD1 Immunotherapy. *Sci. Rep.* **2021**, *11*, 18795. [[CrossRef](#)]
11. Amin, M.B.; Edge, S.; Greene, F.; Byrd, D.R.; Brookland, R.K.; Washington, M.K.; Gershenwald, J.E.; Compton, C.C. *AJCC Cancer Staging Manual*; Springer International Publishing: Berlin/Heidelberg, Germany, 2017; pp. 325–344. [[CrossRef](#)]

12. Hatt, M.; Majdoub, M.; Vallières, M.; Tixier, F.; Le Rest, C.C.; Groheux, D.; Hindie, E.; Martineau, A.; Pradier, O.; Hustinx, R.; et al. 18F-FDG PET Uptake Characterization through Texture Analysis: Investigating the Complementary Nature of Heterogeneity and Functional Tumor Volume in a Multi-Cancer Site Patient Cohort. *J. Nucl. Med.* **2015**, *56*, 38–44. [[CrossRef](#)]
13. Nioche, C.; Orhac, F.; Boughdad, S.; Reuzé, S.; Goya-Outi, J.; Robert, C.; Pellot-Barakat, C.; Soussan, M.; Frouin, F.; Buvat, I. 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. [[CrossRef](#)] [[PubMed](#)]
14. Thie, J.A. Understanding the Standardized Uptake Value, Its Methods, and Implications for Usage. *J. Nucl. Med.* **2004**, *45*, 1431–1434. [[PubMed](#)]
15. Zwanenburg, A.; Vallières, M.; Abdalah, M.A.; Aerts, H.J.W.L.; Andrearczyk, V.; Apte, A.; Ashrafinia, S.; Bakas, S.; Beukinga, R.J.; Boellaard, R.; et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-Based Phenotyping. *Radiology* **2020**, *295*, 328–338. [[CrossRef](#)]
16. Bashir, U.; Azad, G.; Siddique, M.M.; Dhillon, S.; Patel, N.; Bassett, P.; Landau, D.; Goh, V.; Cook, G. The Effects of Segmentation Algorithms on the Measurement of (18)F-FDG PET Texture Parameters in Non-Small Cell Lung Cancer. *EJNMMI Res.* **2017**, *7*, 60. [[CrossRef](#)]
17. Wahl, R.L.; Jacene, H.; Kasamon, Y.; Lodge, M.A. From RECIST to PERCIST: Evolving Considerations for PET Response Criteria in Solid Tumors. *J. Nucl. Med.* **2009**, *50*, 122S–150S. [[CrossRef](#)] [[PubMed](#)]
18. Orhac, F.; Soussan, M.; Chouahnia, K.; Martinod, E.; Buvat, I. 18F-FDG PET-Derived Textural Indices Reflect Tissue-Specific Uptake Pattern in Non-Small Cell Lung Cancer. *PLoS ONE* **2015**, *10*, e0145063. [[CrossRef](#)]
19. Johnson, W.E.; Li, C.; Rabinovic, A. Adjusting Batch Effects in Microarray Expression Data Using Empirical Bayes Methods. *Biostatistics* **2007**, *8*, 118–127. [[CrossRef](#)] [[PubMed](#)]
20. Orhac, F.; Boughdad, S.; Philippe, C.; Stalla-Bourdillon, H.; Nioche, C.; Champion, L.; Soussan, M.; Frouin, F.; Frouin, V.; Buvat, I. A Postreconstruction Harmonization Method for Multicenter Radiomic Studies in PET. *J. Nucl. Med.* **2018**, *59*, 1321–1328. [[CrossRef](#)] [[PubMed](#)]
21. Parmar, C.; Grossmann, P.; Bussink, J.; Lambin, P.; Aerts, H.J.W.L. Machine Learning Methods for Quantitative Radiomic Biomarkers. *Sci. Rep.* **2015**, *5*, 13087. [[CrossRef](#)]
22. Hatt, M.; Tixier, F.; Pierce, L.; Kinahan, P.E.; Rest, C.C.L.; Visvikis, D. Characterization of PET/CT Images Using Texture Analysis: The Past, the Present . . . Any Future? *Eur. J. Nucl. Med. Mol. Imaging* **2017**, *44*, 151–165. [[CrossRef](#)]
23. Oubel, E.; Beaumont, H.; Iannessi, A. Mutual Information-Based Feature Selection for Radiomics. In *Medical Imaging 2016: PACS and Imaging Informatics: Next Generation and Innovations*; SPIE Digital Library: Bellingham, WA, USA, 2016; Volume 9789, pp. 109–117.
24. Xu, Y.; Liu, J.; Hu, Q.; Chen, Z.; Du, X.; Heng, P.A. F-Score Feature Selection Method May Improve Texture-Based Liver Segmentation Strategies. In Proceedings of the 2008 IEEE International Symposium on IT in Medicine and Education, Xiamen, China, 12–14 December 2008; pp. 697–702. [[CrossRef](#)]
25. Ahn, H.K.; Lee, H.; Kim, S.G.; Hyun, S.H. Pre-Treatment 18F-FDG PET-Based Radiomics Predict Survival in Resected Non-Small Cell Lung Cancer. *Clin. Radiol.* **2019**, *74*, 467–473. [[CrossRef](#)] [[PubMed](#)]
26. Xie, C.; Du, R.; Ho, J.W.; Pang, H.H.; Chiu, K.W.; Lee, E.Y.; Vardhanabhuti, V. Effect of Machine Learning Re-Sampling Techniques for Imbalanced Datasets in 18F-FDG PET-Based Radiomics Model on Prognostication Performance in Cohorts of Head and Neck Cancer Patients. *Eur. J. Nucl. Med. Mol. Imaging* **2020**, *47*, 2826–2835. [[CrossRef](#)] [[PubMed](#)]
27. Fawcett, T. An Introduction to ROC Analysis. *Pattern Recognit. Lett.* **2006**, *27*, 861–874. [[CrossRef](#)]
28. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2019.
29. Dittrich, D.; Pyka, T.; Scheidhauer, K.; Lütje, S.; Essler, M.; Bundschuh, R.A. Textural features in FGD-PET/CT can predict outcome in melanoma patients to treatment with Vemurafenib and Ipililumab. *Nuklearmedizin* **2020**, *59*, 228–234. [[CrossRef](#)] [[PubMed](#)]
30. Shaikhina, T.; Lowe, D.; Daga, S.; Briggs, D.; Higgins, R.; Khovanova, N. Decision Tree and Random Forest Models for Outcome Prediction in Antibody Incompatible Kidney Transplantation. *Biomed. Signal Process. Control.* **2019**, *52*, 456–462. [[CrossRef](#)]
31. Wang, M.; Feng, Z.; Zhou, L.; Zhang, L.; Hao, X.; Zhai, J. Computed-Tomography-Based Radiomics Model for Predicting the Malignant Potential of Gastrointestinal Stromal Tumors Preoperatively: A Multi-Classifer and Multicenter Study. *Front. Oncol.* **2021**, *11*, 582847. [[CrossRef](#)]
32. Hotta, M.; Minamimoto, R.; Miwa, K. 11C-Methionine-PET for Differentiating Recurrent Brain Tumor from Radiation Necrosis: Radiomics Approach with Random Forest Classifier. *Sci. Rep.* **2019**, *9*, 15666. [[CrossRef](#)]
33. He, B.; Zhao, W.; Pi, J.-Y.; Han, D.; Jiang, Y.-M.; Zhang, Z.-G.; Zhao, W. A Biomarker Basing on Radiomics for the Prediction of Overall Survival in Non-Small Cell Lung Cancer Patients. *Respir. Res.* **2018**, *19*, 199. [[CrossRef](#)]
34. Seymour, L.; Bogaerts, J.; Perrone, A.; Ford, R.; Schwartz, L.H.; Mandrekar, S.; Lin, N.U.; Litière, S.; Dancey, J.; Chen, A.; et al. IRECIST: Guidelines for Response Criteria for Use in Trials Testing Immunotherapeutics. *Lancet Oncol.* **2017**, *18*, e143–e152. [[CrossRef](#)]
35. Mulkey, F.; Theoret, M.R.; Keegan, P.; Pazdur, R.; Sridhara, R. Comparison of IRECIST versus RECIST V.1.1 in Patients Treated with an Anti-PD-1 or PD-L1 Antibody: Pooled FDA Analysis. *J. Immunother. Cancer* **2020**, *8*, e000146. [[CrossRef](#)]



36. Galavis, P.E.; Hollensen, C.; Jallow, N.; Paliwal, B.; Jeraj, R. Variability of Textural Features in FDG PET Images Due to Different Acquisition Modes and Reconstruction Parameters. *Acta Oncol.* **2010**, *49*, 1012–1016. [[CrossRef](#)] [[PubMed](#)]
37. Dissaux, G.; Visvikis, D.; Da-ano, R.; Pradier, O.; Chajon, E.; Barillot, I.; Duverge, L.; Masson, I.; Abgral, R.; Santiago Ribeiro, M.-J.; et al. Pre-Treatment 18F-FDG PET/CT Radiomics Predict Local Recurrence in Patients Treated with Stereotactic Radiotherapy for Early-Stage Non-Small Cell Lung Cancer: A Multicentric Study. *J. Nucl. Med.* **2020**, *61*, 814–820. [[CrossRef](#)] [[PubMed](#)]
38. Nakajo, M.; Jinguji, M.; Tani, A.; Kikuno, H.; Hirahara, D.; Togami, S.; Kobayashi, H.; Yoshiura, T. Application of a Machine Learning Approach for the Analysis of Clinical and Radiomic Features of Pretreatment [18F]-FDG PET/CT to Predict Prognosis of Patients with Endometrial Cancer. *Mol. Imaging Biol.* **2021**, *23*, 756–765. [[CrossRef](#)] [[PubMed](#)]