

Article

Prediction of the Presence of Targetable Molecular Alteration(s) with Clinico-Metabolic 18 F-FDG PET Radiomics in Non-Asian Lung Adenocarcinoma Patients

Nicolas Aide ¹ , Kathleen Weys ² and Charline Lasnon ^{1,2,*} ¹ UNICAEN, INSERM 1086 ANTICIPE, Normandy University, 14000 Caen, France² Nuclear Medicine Department, Comprehensive Cancer Centre F. Baclesse, UNICANCER, 14000 Caen, France

* Correspondence: c.lasnon@baclesse.unicancer.fr; Tel.: +33-261-455-268; Fax: +33-231-455-101

Abstract: This study aimed to investigate if combining clinical characteristics with pre-therapeutic 18 F-fluorodeoxyglucose (18 F-FDG) positron emission tomography (PET) radiomics could predict the presence of molecular alteration(s) in key molecular targets in lung adenocarcinoma. This non-interventional monocentric study included patients with newly diagnosed lung adenocarcinoma referred for baseline PET who had tumour molecular analyses. The data were randomly split into training and test datasets. LASSO regression with 100-fold cross-validation was performed, including sex, age, smoking history, AJCC cancer stage and 31 PET variables. In total, 109 patients were analysed, and it was found that 63 (57.8%) patients had at least one molecular alteration. Using the training dataset ($n = 87$), the model included 10 variables, namely age, sex, smoking history, AJCC stage, excessKustosis_HISTO, sphericity_SHAPE, variance_GLCM, correlation_GLCM, LZE_GLZLM, and GLNU_GLZLM. The ROC analysis for molecular alteration prediction using this model found an AUC equal to 0.866 ($p < 0.0001$). A cut-off value set to 0.48 led to a sensitivity of 90.6% and a positive likelihood ratio (LR+) value equal to 2.4. After application of this cut-off value in the unseen test dataset of patients ($n = 22$), the test presented a sensitivity equal to 90.0% and an LR+ value of 1.35. A clinico-metabolic 18 F-FDG PET phenotype allows the detection of key molecular target alterations with high sensitivity and negative predictive value. Hence, it opens the way to the selection of patients for molecular analysis.

Keywords: adenocarcinoma; FDG; lung cancer; molecular analysis; pet; radiomic



Citation: Aide, N.; Weys, K.; Lasnon, C. Prediction of the Presence of Targetable Molecular Alteration(s) with Clinico-Metabolic 18 F-FDG PET Radiomics in Non-Asian Lung Adenocarcinoma Patients. *Diagnostics* **2022**, *12*, 2448. <https://doi.org/10.3390/diagnostics12102448>

Academic Editor: Anna Pecorelli

Received: 12 September 2022

Accepted: 7 October 2022

Published: 10 October 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/>).

1. Introduction

Lung cancer is the leading cause of cancer death in France as well as the world [1]. It is also one of the main worldwide indications of 18 F-fluorodeoxyglucose (18 F-FDG) positron emission tomography (PET) in nuclear medicine departments [2] and a research topic of major interest. Recently, many publications have explored the association of tumour mutations, especially epidermal growth factor receptor (EGFR), anaplastic lymphoma kinase (ALK), Kirsten rat sarcoma viral oncogene homolog (KRAS), and ROS1 mutations, with 18 F-FDG radiomics features for in vivo non-invasive diagnostic or prognostic purposes [3–5].

Meanwhile, in oncology genomics, next-generation sequencing (NGS) that allows the detection of multiple anomalies on different genomics scales, especially mutations and structural variations, supplanted traditional methods based on multiple tests. The strength of NGS is its ability to be performed with a small amount of tissue from a single biopsy [6], a single extraction, and a single test, which is both time- and cost-effective [7–9]. Although the whole genome sequencing remains a gold standard in oncology genomics, large DNA-based NGS selected panels conducted on cancer tissue samples of patients (and maybe in the near future on circulating cell-free DNA) have arisen as convenient alternatives [10].

These panels aim at identifying targetable molecular alterations helpful for the personalised treatment of patients. In non-small cell lung cancer (NSCLC), and specifically in lung adenocarcinoma, the NGS is increasingly carried out at diagnosis, since the first line of treatment depends on multiple molecular targets [11], but no consensus has been reached until now. The current hot topic is to determine who will benefit from NGS panels and when in the care timeline.

In this regard, the present study aimed to investigate the combination of usual clinical characteristics (i.e., sex, age, smoking history, cancer stage according to the American Joint Committee on Cancer (AJCC)) and pre-therapeutic 18 F-FDG PET radiomics. The purpose was to find out if it could predict the existence of molecular alterations in key molecular targets in lung adenocarcinoma (namely EGFR, v-raf murine sarcoma viral oncogene homolog B1 (BRAF), KRAS, neuroblastoma RAS viral oncogene homolog (NRAS), mesenchymal-epithelial transition (MET), serine/threonine kinase 11 (STK11), phosphatidylinositol 3-kinase catalytic alpha polypeptide (PIK3CA), ALK and ROS1) to screen patients who are more likely to benefit from molecular testing.

2. Materials and Methods

2.1. Population

This non-interventional monocentric study had a retrospective design. The study population included all patients with newly diagnosed lung adenocarcinoma who had an NGS exploration of panel Colon and Lung Cancer Panel v3 (CLv3). The subjects were selected from those patients who referred to our PET unit for their initial staging between January 2018 and December 2019. Their sex, age, smoking history, and AJCC stage were recorded. The institutional review board approved the study and waived the need for informed signed consent. Following the European General Data Protection Regulation, data collection for this study was approved by the national committee for data privacy, with the registration code of N° F20210309115801.

2.2. Immunohistochemistry

Automated immunohistochemistry using a Ventana Bench Mark Ultra was performed on 4 μ m-thick paraffin sections of biopsy with clone D4D6 for ROS1 (pre-diluted). A complementary fluorescence in situ hybridisation (FISH) was systematically performed in case of positivity [12].

2.3. Next-Generation Sequencing Panel Colon and Lung Cancer Panel v3 Analysis [13]

Analyses were carried out using a fixed and included paraffin sample. The tumour genomic DNA extraction was conducted with the Q1Aamp DNA FFPE Tissue Kit on Q1Acube (Q1AGEN). The NGS was performed using an Ion Personal Genome Machine (Life Technologies, Villebon sur Yvette, France). The average depth was >500 X; on target $>90\%$. Bioinformatic analyses (alignment and call of variants and annotations):

- Life Technologies: Torrent Suite (version 5.6), Variant Caller (version 5.6), Ion Reporter (version 5.6)
- Nextgene (version 2.4.1.2, Softgenetics, State College, PA, USA).

The copy number variant analysis was expressed as the ratio of mean depths to amplicons ± 2 . The detection limit was set to 3% for punctual mutation and 5% for insertion/deletion for a minimum depth of 100X per amplicon. Variations of sequences recognised as non-pathogenic were not mentioned.

2.4. Positron Emission Tomography Acquisition and Analysis

Patient care and 18 F-FDG administration (3.0 MBq/kg) were performed according to the guidelines of the European Association of Nuclear Medicine for oncologic examinations [14]. The PET acquisitions were acquired on 2 systems:

- (1) A TrueV analogic PET/CT (Siemens Healthineers, Erlangen, Germany) with three iterations and 21 subsets with point spread function (PSF) reconstruction ($2.0 \times 4.0 \times 4.0 \text{ mm}^3$ voxels). The PET emission acquisition was performed from skull to mid-thighs for 2 min and 40 s and 3 min and 40 s per bed position for normal-weight and overweight patients, respectively.
- (2) A Vereos digital PET/CT (Philips Medical Solutions, USA) with two iterations and 10 subsets with PSF reconstruction (2 mm^3 voxels). The PET emission acquisition was performed from the skull to mid-thighs for 2 min per bed position regardless of the body habitus of the patients.

An experienced nuclear physician drew volumes of interest over the hypermetabolism of the primary lung lesion using a gradient-based delineation, showing to outperform threshold-based methods in terms of accuracy and robustness [15] (PET edge) on MIM software (version 5.6.5, MIM Software Inc., Cleveland, OH, USA) and recorded as RTstruct files. Afterward, the RTstruct files were uploaded in LifeX 6.3 software [16], and the automatic close function (3D dilatation followed by erosion of 10 voxels) was systematically applied to account for any hypo-metabolic area(s), such as necrotic parts of the tumour volume. No other changes were made, especially freehand modifications. A small lesion size leading to a volume of interest containing less than 64 voxels was an exclusion criterion according to the LifeX software procedural standards. Indeed, this is a well-known impairment for the calculation of metabolic ^{18}F -FDG-PET heterogeneity [17].

The following parameters, fulfilling the Image Biomarker Standardization Initiative [18], were extracted from PET images using an absolute resampling of 64 grey levels and standardised uptake value (SUV) comprised between 0 and 30:

- Conventional parameters: SUV_{mean} , SUV_{max} , metabolic tumour volume, and total lesion glycolysis
- Histogram parameters: skewness_HISTO, kurtosis_HISTO, excessKurtosis_HISTO, entropy_log2_HISTO, and uniformity_HISTO
- Shape parameters: sphericity_SHAPE and compacity_SHAPE
- Grey-Level Co-Occurrence Matrix (GLCM) parameters: inverse difference_GLCM, angular second moment_GLCM, variance_GLCM, correlation_GLCM, joint entropy_GLCM, and dissimilarity_GLCM
- Neighbouring grey-level dependence matrix (NGLDM) parameters: coarseness_NGLDM, contrast_NGLDM, and busyness_NGLDM
- Grey-level zone length matrix (GLZLM) parameters: SZE_GLZLM, LZE_GLZLM, LGZE_GLZLM, HGZE_GLZLM, SZLGE_GLZLM, SZHGE_GLZLM, LZLGE_GLZLM, LZHGE_GLZLM, GLNU_GLZLM, ZLNU_GLZLM, and ZP_GLZLM

Detailed specifications for these parameters are available on the Image Biomarker Standards Initiative website [19].

A post-reconstruction harmonisation was run using ComBaT for all PET parameters found to be statistically significantly different between PET systems [20].

2.5. Statistical Analysis

The data are presented as mean values unless otherwise specified. The PET parameters extracted from the 2 PET systems described above were compared using a Mann–Whitney non-parametric test before and after ComBat harmonisation applying Bonferroni correction. A process was used by randomly splitting the data in training (87 patients, 80%) and test (22 patients, 20%) datasets. Instead of creating a separate validation dataset, a k-fold cross-validation was used with the training dataset to tune model parameters.

To select features of interest and build a model of prediction, the LASSO regression with a binary response variable (0: no molecular alteration, 1: at least one molecular alteration) was performed on the training dataset of patients with 100-fold cross-validation, including the following explanatory variables: sex, age, smoking history, AJCC stage and all the previously described PET variables. The optimal pred.core cut-off value was determined using the receiver operating characteristic curve. This pred.score cut-off value

was then tested on the unseen test dataset. Sensitivity (also known as recall), specificity, positive predictive value (also known as precision), negative predictive value, positive likelihood ratio (LR+), negative likelihood ratio (LR−), and accuracy were reported. LRs are defined as follows:

$$LR+ = \frac{\text{sensitivity}}{1 - \text{specificity}} \quad (1)$$

$$LR- = \frac{1 - \text{sensitivity}}{\text{specificity}} \quad (2)$$

The LR includes the sensitivity and specificity of a test into a single measure. The best diagnostic test to detect the disease is the one with the larger LR+ [21,22].

Statistical analysis and graphs were made using XLSTAT software (version 2019: Data Analysis and Statistical Solution for Microsoft Excel. Addinsoft). A *p*-value of less than 0.05 was considered statistically significant unless otherwise specified.

3. Results

3.1. Patients and Next-Generation Sequencing Characteristics of the Entire Data Population

In total, 111 patients were screened from January 2018 to December 2019. Two patients were excluded due to the small volume of their lesions (3.0 and 1.4 mL). Finally, 109 patients were included. The database was composed of 34 female and 75 male subjects (sex ratio M/F = 2.2) with a median age of 66 years (range: 37–86). Moreover, 96 patients (88.1%) had a history of smoking. There were 19 (17.4%), 37 (34.0%), and 53 (48.6%) patients at AJCC stages of I or II, III, and IV, respectively.

In addition, 63 patients (57.8%) had at least one molecular alteration. Among them, 53, 8, and 2 patients had one (84.1%), two (12.7%), and three alterations (3.2%), respectively. Figure 1 displays a detailed description. With regard to EGFR mutations, nine, three, and two were in exon19, exon20, and exon21, respectively. The KRAS mutations were all in the exon2 except for one patient. It is noteworthy that no HER2 mutation occurred.

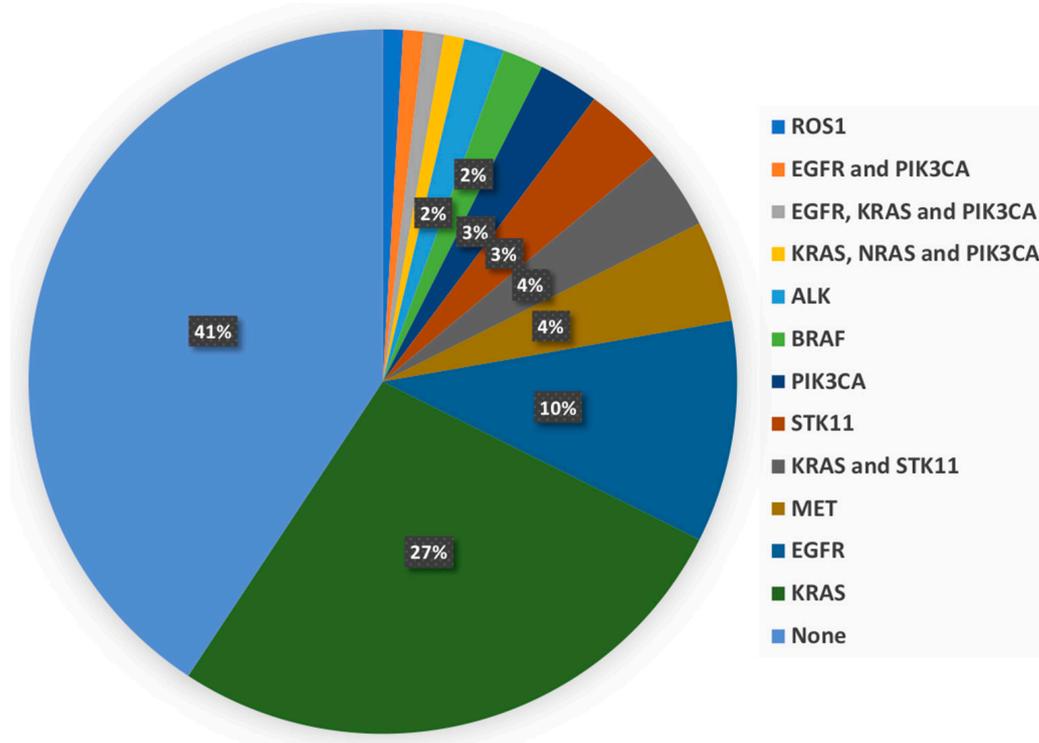


Figure 1. Molecular alterations description of patients.

The median delay between the completion of the biopsy and the availability of the results for all previously described mutations was 23 days. For 100 patients (91.7%), the PET/CT examination was performed before the availability of the genetic results with a median interval of 39 days. For 69 patients (63.3%), the PET/CT was even performed before the biopsy with a median interval of 21 days. Table 1 summarises the clinical characteristics of the training and test datasets of patients.

Table 1. Molecular and clinical characteristics of patients.

Variable	Categories	Training Group (n = 87)		Test Group (n = 22)	
		Frequencies	%	Frequencies	%
Molecular alterations	None	34	39.1	12	54.5
	At least one	53	60.9	10	45.5
Gender	Female	25	28.7	9	40.9
	Male	62	71.3	13	59.1
Smoking history	No	8	9.2	5	22.7
	Yes	79	90.8	17	77.3
AJCC stage	I or II	17	19.5	2	9.1
	III	28	32.2	9	40.9
	IV	42	48.3	11	50.0

3.2. Positron Emission Tomography Data Harmonisation

In total, 51 patients (46.8%) underwent their examinations on the TrueV analogic system and 58 (53.2%) underwent their examinations on the Vereos digital system. None of the conventional and histogram PET parameters were significantly different between PET systems (Supplemental Table S1). The parameters found to be statistically different among PET systems were compacity_SHAPE, variance_GLCM, correlation_GLCM, dissimilarity_GLCM, coarseness_NGLDM, contrast_NGLDM, SZE_GLZLM, LZE_GLZLM, LGZE_GLZLM, SZLGE_GLZLM, LZHE_GLZLM, and ZP_GLZLM. However, after ComBat harmonisation, only LZHE_GLZLM remained different among PET systems ($p < 0.0001$, Table 2); therefore, this parameter was not further considered.

Table 2. Comparison of TrueV and Vereos positron emission tomography quantitative variables before and after the ComBat harmonisation process. The only variables found to be significantly different between the TrueV and the Vereos systems before the ComBat harmonisation process are presented here. To see the data for all variables, please refer to the exhaustive supplemental Table S1 (* according to Bonferroni correction, a p -value < 0.0016 was considered significant).

Variable	Min	Max	Mean	p Value *	Min	Max	Mean	p Value *
Shape parameters								
Compacity TrueV	0.735	8.699	2.401		0.492	7.707	3.058	
Compacity Vereos	1.311	11.225	4.439	<0.0001	1.027	7.727	3.048	0.963
GLCM parameters								
Variance TrueV	2.169	207.117	28.928		0.383	145.670	19.352	
Variance Vereos	1.019	60.837	10.932	<0.0001	−4.923	141.562	19.352	0.854
Correlation TrueV	0.211	0.789	0.651		0.278	0.861	0.721	
Correlation Vereos	0.356	0.945	0.784	<0.0001	0.305	0.879	0.721	0.892
Dissimilarity TrueV	0.952	10.475	3.422		0.839	8.457	2.815	
Dissimilarity Vereos	0.715	6411	2.281	0.0002	0.685	8.429	2.815	0.839
NGLDM parameters								
Coarseness TrueV	0.001	0.070	0.018		−0.001	0.057	0.014	
Coarseness Vereos	0.000	0.054	0.009	0.001	0.002	0.070	0.014	0.674
Contrast TrueV	0.014	0.878	0.198		0.000	0.664	0.141	
Contrast Vereos	0.008	0.618	0.091	<0.0001	0.014	0.942	0.141	0.774

Table 2. Cont.

Variable	Min	Max	Mean	p Value *	Min	Max	Mean	p Value *
GLZLM parameters								
SZE TrueV	0.229	0.775	0.536		0.211	0.723	0.499	
SZE Vereos	0.230	0.704	0.466	0.0004	0.252	0.748	0.499	0.802
LZE TrueV	4.368	18,386.635	1127.261		-13,631.716	635,105.336	25,996.822	
LZE Vereos	3.600	1,191,735.375	47,864.883	<0.0001	-8620.105	853,331.114	25,996.822	0.0002
LGZE TrueV	0.004	0.225	0.034		0.002	0.166	0.024	
LGZE Vereos	0.001	0.087	0.016	<0.0001	-0.001	0.148	0.024	0.631
SZLGE TrueV	0.003	0.103	0.013		0.002	0.076	0.010	
SZLGE Vereos	0.001	0.045	0.007	<0.0001	-0.001	0.080	0.010	0.353
LZHGE TrueV	666.714	1,258,030.400	50,296.736		-155,560.617	43,339,050.699	156,236.634	
LZHGE Vereos	1793.265	55,736,935.663	2,889,821.717	<0.0001	-527,621.307	39,784,583.618	1,561,236.634	<0.0001
ZP TrueV	0.033	0.616	0.255		-0.014	0.537	0.196	
ZP Vereos	0.007	0.625	0.144	<0.0001	0.054	0.695	0.196	0.839

3.3. Construction of Prediction Model Using a Lasso Regression with a Cross-Validation on the Training Dataset (n = 87)

The model included 10 variables, namely age, sex, smoking history, AJCC stage, excessKurtosis_HISTO, sphericity_SHAPE, variance_GLCM, correlation_GLCM, LZE_GLZLM, and GLNU_GLZLM (Figures 2 and 3). The optimal lambda determined by cross-validation was equal to 0.030, and the corresponding coefficients can be seen in Supplemental Table S2.

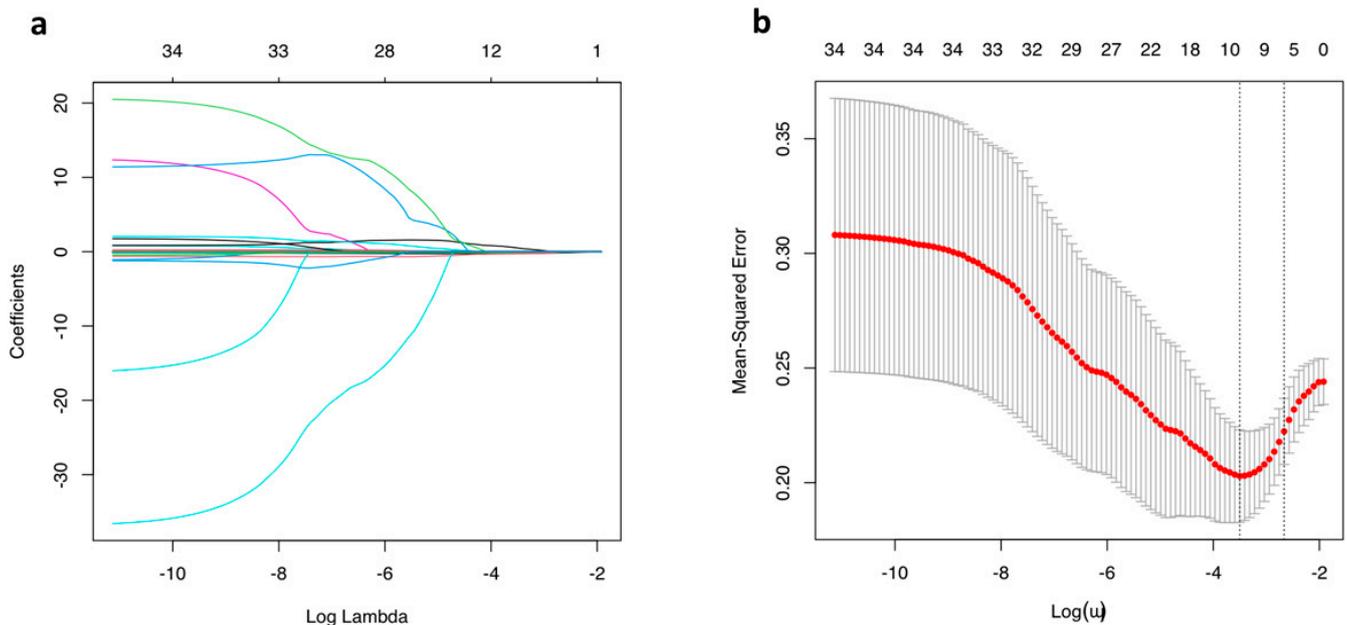


Figure 2. Coefficients (a) and cross-validation (b) plots of the LASSO analysis.

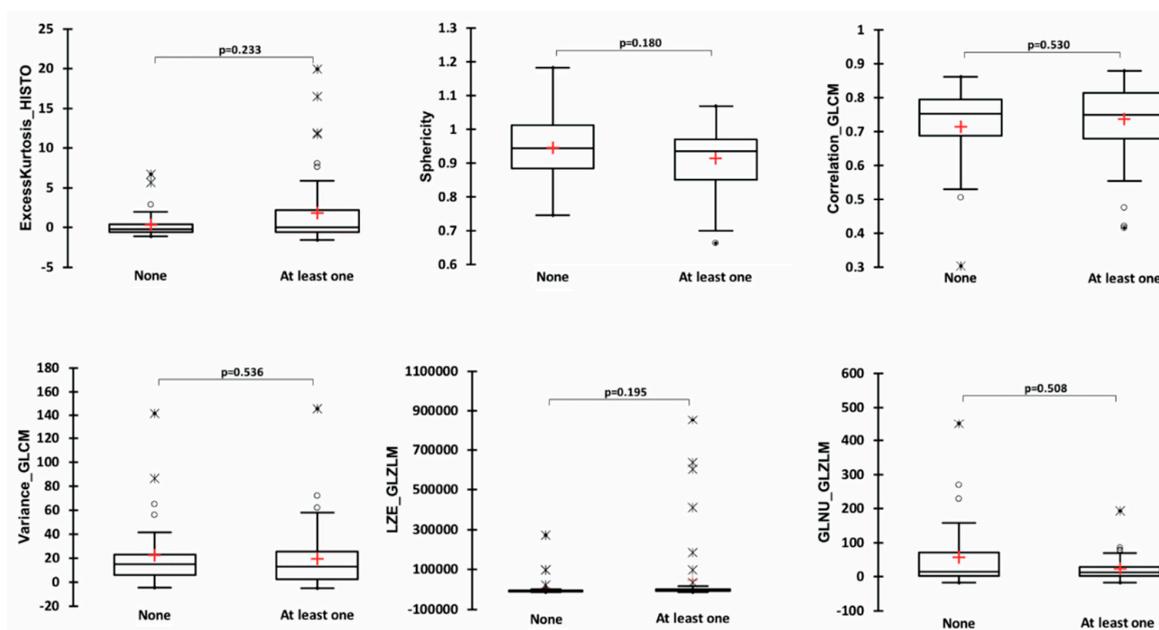


Figure 3. Comparison of patients without any molecular alteration and patients with at least one in terms of all positron emission tomography quantitative variables included in the model.

3.4. Relationship between Variables Included in the Lasso Regression Model

Significant correlations were observed among PET variables of interest, which are summarised in detail in Table 3. The strongest correlation was observed between variance_GLCM and LZE_GLZLM with $\rho = -0.702$ and $p < 0.0001$. There was no correlation between age and PET variables of interest (Table 3). Moreover, there were no differences in PET variables of interest in terms of sex, smoking history, and AJCC stages (Table 4).

Table 3. Spearman correlations matrix for positron emission tomography variables of interest and age. Values in bold are those that correspond to statistically significant results.

Variables	Age	ExcessKurtosis_HISTO	Sphericity_SHAPE	Variance_GLCM	Correlation_GLCM	LZE_GLZLM	GLNU_GLZLM
Age	1	-0.176	0.059	-0.105	0.067	0.169	0.083
ExcessKurtosis_HISTO	-0.176	1	-0.344	-0.369	-0.075	0.401	-0.232
Sphericity_SHAPE	0.059	-0.344	1	0.212	-0.318	-0.307	-0.367
Variance_GLCM	-0.105	-0.369	0.212	1	-0.153	-0.702	0.167
Correlation_GLCM	0.067	-0.075	-0.318	-0.153	1	0.362	0.574
LZE_GLZLM	0.169	0.401	-0.307	-0.702	0.362	1	0.263
GLNU_GLZLM	0.083	-0.232	-0.367	0.167	0.574	0.263	1

Table 4. Comparisons of positron emission tomography variables of interest according to sex, smoking history, and stage according to American Joint Committee on Cancer.

Variables		ExcessKurtosis_HISTO	Sphericity_SHAPE	Variance_GLCM	Correlation_GLCM	LZE_GLZLM	GLNU_GLZLM
Sex, mean (SD)	Females (n = 25)	1.849 (4.830)	0.904 (0.096)	20.010 (30.118)	0.743 (0.104)	35343.472 (133495.181)	19.109 (23.159)
	Males (n = 62)	0.967 (3.213)	0.936 (0.094)	20.979 (25.708)	0.722 (0.113)	30955.319 (145426.809)	44.254 (80.101)
	p value	0.685	0.140	0.611	0.453	0.476	0.748
Smoking history, mean (SD)	No (n = 8)	-0.270 (0.653)	0.930 (0.078)	19.039 (15.733)	0.744 (0.083)	-4966.920 (9276.960)	26.023 (27.943)
	Yes (n = 79)	1.372 (3.891)	0.927 (0.097)	20.869 (27.814)	0.726 (0.113)	35981.670 (147862.953)	38.143 (72.412)
	p value	0.260	0.891	0.670	0.903	0.608	0.812
AJCC stage, mean (SD)	I (n = 5)	4.180 (7.410)	0.969 (0.090)	14.265 (22.681)	0.574 (0.151)	-6434.922 (2239.339)	2.494 (1.497)
	II (n = 12)	1.271 (3.842)	0.886 (0.083)	15.924 (15.299)	0.755 (0.095)	-6371.951 (8811.401)	44.015 (57.732)
	III (n = 28)	0.426 (2.868)	0.950 (0.084)	21.956 (30.266)	0.736 (0.120)	29371.230 (119,758.438)	41.815 (62.151)
	IV (n = 42)	1.384 (3.619)	0.918 (0.102)	21.995 (27.999)	0.733 (0.091)	49,739.528 (17,7486.562)	35.952 (80.746)
	p value	0.092	0.186	0.862	0.066	0.329	0.195

3.5. Comparison of Variables Included in the Lasso Regression Model between Patients with and without Molecular Alteration(s)

There was no significant difference between patients without any molecular alterations and patients with at least one molecular alteration in terms of age, excessKustosis_HISTO, sphericity_SHAPE, variance_GLCM, correlation_GLCM, LZE_GLZLM, and GLNU_GLZLM (Figure 3). There was no association between molecular status and AJCC stage ($p = 0.102$, Figure 4).

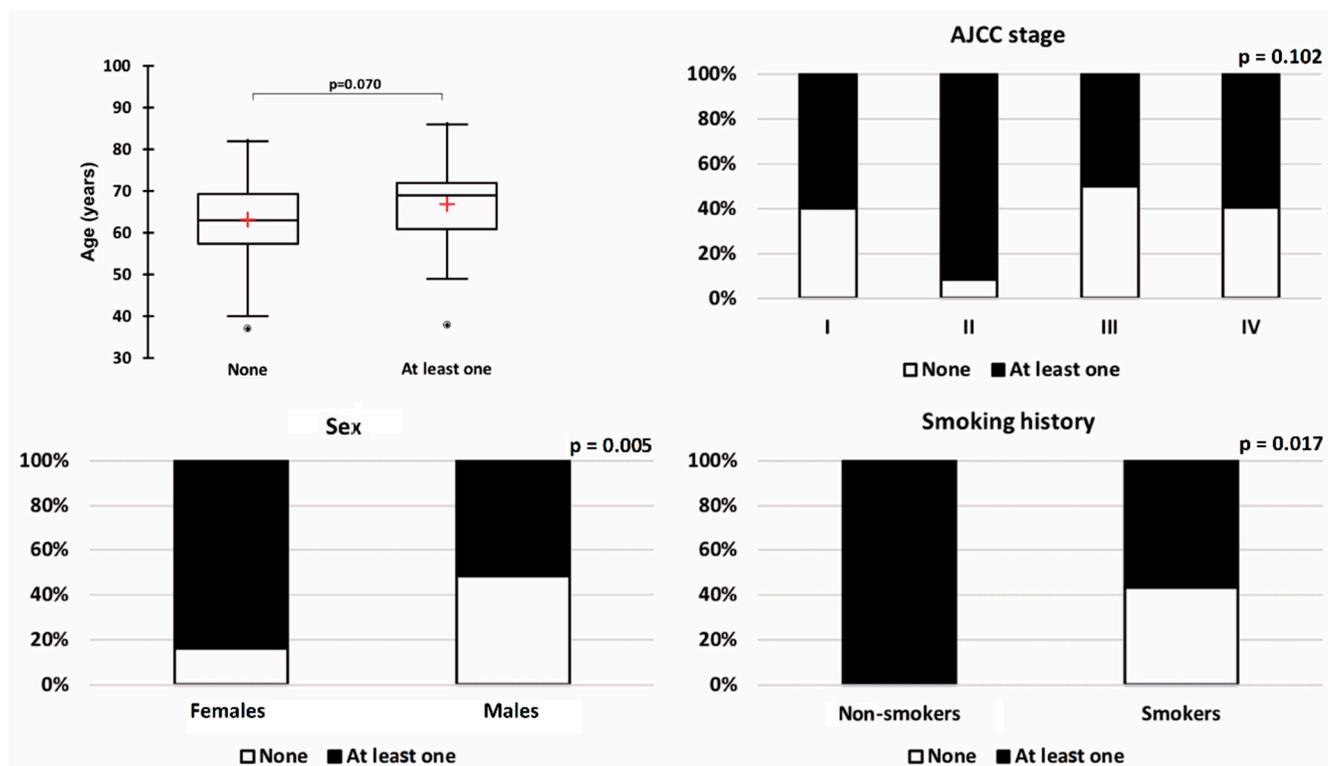


Figure 4. Comparison of patients without any molecular alteration and patients with at least one for all the clinical variables included in the model.

Sex and smoking history were significantly associated with the molecular status. More molecular alterations were observed in females (84.0%) compared to males (51.6%) ($p = 0.005$, Figure 3). The KRAS and EGFR mutations represented the great majority of molecular alterations in females, as 56.0% and 28.0% of them had KRAS and EGFR mutations, respectively. In addition, the only positive ROS1 translocation-related protein expression was observed in a woman. A complementary FISH was performed confirming the presence of an ROS1 translocation in 80% of the tumour cells visualised. In addition, in the training dataset, all non-smokers ($n = 8$) presented at least one molecular alteration ($p = 0.017$).

3.6. LASSO Regression Model Diagnostic Performances for Molecular Alteration(s) Detection in the Training Dataset ($n = 87$)

The receiver operating characteristic analysis for molecular alteration prediction using this LASSO regression model in the training dataset of patients found an area under the curve equal to 0.866 (95%CI = 0.792–0.941, $p < 0.0001$). The pred.score cut-off value (or operating point) was chosen by optimising utility with the aim of optimising the sensitivity of the test to lower the rate of false-negative cases (FN) as appropriate for a screening test. A cut-off value set to 0.48 led to a sensitivity of 90.6%, a negative predictive value of 80.8%, and an LR+ of 2.37 (specificity = 61.8%, positive predictive value = 78.7%, accuracy of 79.3%). It should be mentioned that in this configuration, 26 patients (29.9%) tested negative. Among them, only five patients had FN results, and none of them had EGFR mutations

(one ALK/KRAS, one KRAS, two STK11, and one PIK3CA). Among the 61 patients tested positive, 49 (80.3%) were true positives. These true-positive patients had a median age of 69 years (range = 38–86), 56.3% were males, and 83.3% had smoking history.

3.7. Prediction Model Screening Performances on the Unseen Test Dataset ($N = 22$)

Applying this cut-off value in the validation dataset of patients, the test presented a sensitivity of 90.0%, a negative predictive value of 80.0%, and an LR+ equal to 1.35 with one FN result among the five negative tests (22.7%) concerning a KRAS mutation. There were nine true-positive patients who had a median age of 66 years (range = 55–83), 22.2% were males, and 77.8% had smoking history. Representative images of a true-positive patient are shown on Figure 5.

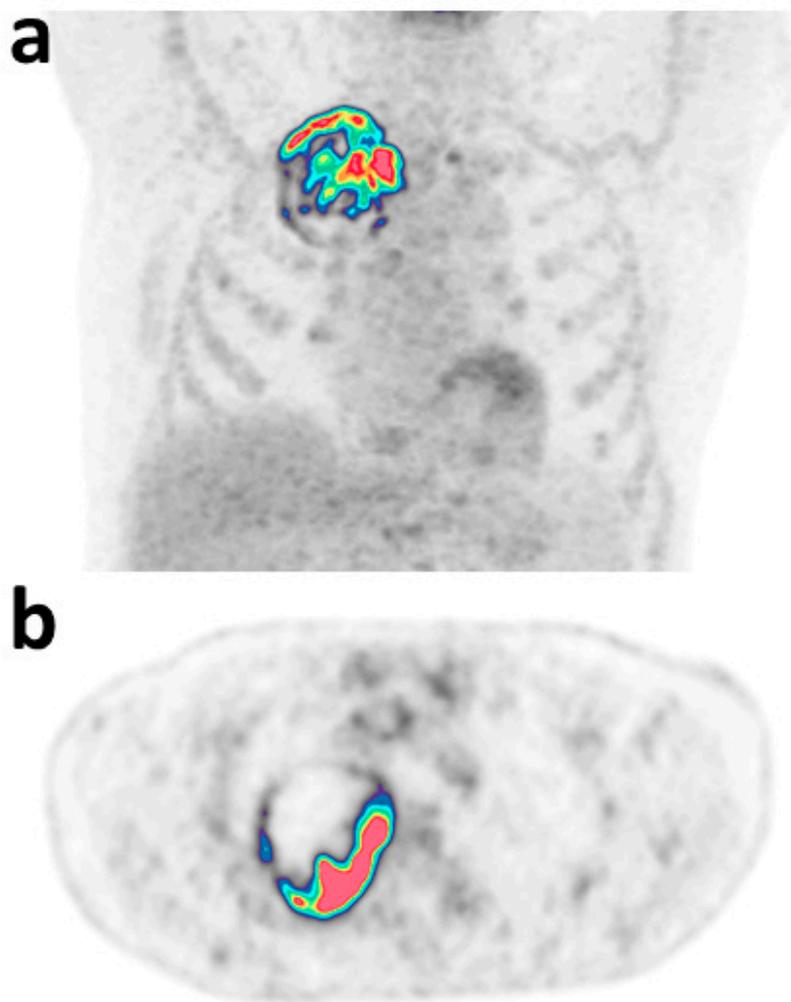


Figure 5. Representative maximum intensity projection (a) and axial (b) images of true-positive patients with MET gene alteration and a pred.score equal to 0.92. Model variables were: age = 70 years, sex = male, smoking history = yes, AJCC stage: IIIA, excessKurtosis_HISTO = 11.72, sphericity_SHAPE = 0.83, variance_GLCM = 24.40, correlation_GLCM = 0.87, LZE_GLZLM = 9188.94, GLNU_GLZLM = 76.81.

4. Discussion

Usage of a model based on clinical and PET radiomics appears to be a promising strategy for screening at the time of diagnosis of lung adenocarcinoma patients who may have a targetable molecular alteration. This could certainly reduce unnecessary costs by avoiding the need to test patients for whom we could know that there is little chance of finding molecular conditions. Moreover, it helps speed the management of these patients

by eliminating the need to wait for the results of an unnecessary test. For instance, in the present study, PET/CT was usually performed more than one month before the availability of genetic analyses results (median = 39 days).

At our institution, somatic mutation detection using immunochemistry and NGS Panel CLv3 and single tests for ROS1 and ALK mutations was performed for nearly all lung adenocarcinoma patients at the time of diagnosis, and finally, almost half of them had negative findings (42%). It is worth noting that frequencies of molecular alterations observed in our database are representative of those previously observed in a Western population with lung cancer [23]. Therefore, it can be said that using the proposed regression model would have avoided 31 molecular tests (26 in the training and five in the test datasets). In other words, applying this strategy could spare almost 1/3 of molecular testing at the cost of six false-negative tests (5.5%).

The included model was a mix of clinical characteristics (age, sex, smoking history, and AJCC stage) and PET characteristics (excessKustosis_HISTO, sphericity_SHAPE, variance_GLCM, correlation_GLCM, LZE_GLZLM, and GLNU_GLZLM). In the sight of the LASSO regression coefficients, the strongest predictive variables were tumour correlation_GLCM, sphericity_SHAPE, sex, and smoking history. Sex and smoking history were significantly different between patients with and without molecular alterations on univariable analysis with an increased risk of molecular alterations in females and non-smokers. This is somewhat concordant with the results of previous studies indicating the higher rate of EGFR mutations [24–26], KRAS mutations [27,28], ROS1 gene fusion [29,30], and STK11 [31] expression in females. Moreover, non-smokers with lung cancer (mostly adenocarcinomas) have more tumour genetic mutations than smokers, according to the results of the sequencing-based studies [32–34]. To date, to predict whether a lung adenocarcinoma might harbour targetable genetic alterations, the clinician considered the aforementioned factors related to the pathobiology of smoke-associated tumours. Thus, it is interesting to note that in our patient dataset, many of the patients with one or more molecular alterations tested positive by the model were male and had a history of smoking. For example, in our training dataset, the true-positive patients were males in 56.3% of cases and had a history of smoking in 83.3% of cases. This supports the fact that the ¹⁸F-FDG/PET signature may have additional value over the already common clinical stratification criteria.

Correlation_GLCM represents the linear dependency of grey levels in GLCM and sphericity is a measure of the roundness of the shape of the tumour region relative to the sphere. It should be noted that correlation_GLCM was among the parameters that needed to be harmonised between the two PET systems. This is in line with the results of previous studies demonstrating the impact of reconstruction parameters on conventional PET metrics and texture features in NSCLC [2,35] and hence, the need for harmonising standards.

Moreover, when it comes to the comparison of analogic and digital PET quantitative variables analysis, it is important to point out that conventional and histogram parameters were not found to be different from our systems in this reconstruction configuration. A harmonisation process was needed only for some second- and third-order textural features and only one failure of the process was observed with LZHGE_GLZLM.

These findings demonstrated for the first time that textural features extracted from digital and analogic PET systems can be pooled using harmonisation strategies currently under development [20,36]. However, it has been shown that applying a smoothing filter with a large kernel as per EARL procedure [37] or using a larger voxel size can lead to the loss of accuracy of radiomics metrics for tumour characterisation purposes [38].

The study had some limitations; first, these encouraging results need to be confirmed by a larger multi-centre study. Furthermore, their extrapolation to other populations, for which the repartition of histological subtypes and mutational status could be different, needs to be investigated [39–41]. Moreover, to ensure its translation into clinical practice, a worldwide harmonisation strategy is needed, and the development of dedicated software for automatic computation of the model equation seems mandatory. Given the flourishing

number of models of this type in the oncology literature for a multitude of hypotheses, this axis of development should be carried out in the short term.

Secondly, since the great emulation in the framework of precision oncology and genetics, our knowledge will surely be in constant evolution, and a strategy that works today will have to be constantly adapted according to new discoveries in the field. Thirdly, the molecular analysis was usually performed on biopsies with the risk of spatial tumour heterogeneity and missing some tumour molecular alterations. For example, Swanton et al. noted that many oncogenic alterations were only identified in specific tumour locations generating tumour heterogeneity [42]. Similarly, Pelosi et al. micro-dissected several tumour regions of different architectures from 20 adenocarcinomas and revealed that 60% of these tumours had intra-tumour molecular heterogeneity [43]. Therefore, it is possible to wonder if some tests considered as false-positive results in our study were not linked to a biopsy sampling error. Indeed, metabolic tumour characterisation considers the entire tumour volume and not just a sample. This can be a strength together with its non-invasiveness.

Finally, spatiotemporal heterogeneity could also be considered. In this study, the strategy was explored at diagnosis, but another time point during patient management could be investigated, as no consensus has yet been reached [44]. The few patients with false-negative tests at diagnosis might benefit from a molecular analysis at another time of their treatment.

In conclusion, screening non-Asian lung adenocarcinoma patients at the time of diagnosis by means of a model including clinical parameters and ¹⁸F-FDG/PET radiomics before performing a tumour molecular analysis seems to be an efficient strategy. It allows predicting the existence of key molecular target(s) with high sensitivity.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/diagnostics12102448/s1>, Table S1: Comparison of TrueV and Vereos positron emission tomography quantitative variables before and after the ComBat harmonization process; Table S2: LASSO regression coefficients. Variables with null coefficients are not displayed.

Author Contributions: Conceptualisation: N.A. and C.L., Data curation: K.W., Formal analysis: C.L., Investigation: C.L., Methodology: C.L., Project administration: C.L., Resources: K.W. Supervision: N.A., Validation: N.A., Visualisation: K.W., Writing—original draft: C.L., Writing—review and editing: N.A. and K.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The authors are accountable for all aspects of the work and ensure that questions related to the accuracy or integrity of any part of the work were appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the institutional review board of the François Baclesse Cancer Center on 12/21/2020 (project identification: FDG-PredGEN1) and declared to the national committee for data privacy, with the registration code: N° F20210309115801.

Informed Consent Statement: Patient consent was waived due to the retrospective design of this study.

Data Availability Statement: All data generated and analysed during this study are included in this published article. Raw data supporting the findings of this study are available from the corresponding author on reasonable request.

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

References

1. de Groot, P.M.; Wu, C.C.; Carter, B.W.; Munden, R.F. The epidemiology of lung cancer. *Transl. Lung Cancer Res.* **2018**, *7*, 220–233. [[CrossRef](#)] [[PubMed](#)]
2. Houdu, B.; Lasnon, C.; Licaj, I.; Thomas, G.; Do, P.; Guizard, A.V.; Desmonds, C.; Aide, N. Why harmonization is needed when using FDG PET/CT as a prognosticator: Demonstration with EARL-compliant SUV as an independent prognostic factor in lung cancer. *Eur. J. Nucl. Med. Mol. Imaging* **2019**, *46*, 421–428. [[CrossRef](#)]

3. Jiang, M.; Zhang, Y.; Xu, J.; Ji, M.; Guo, Y.; Guo, Y.; Xiao, J.; Yao, X.; Shi, H.; Zeng, M. Assessing EGFR gene mutation status in non-small cell lung cancer with imaging features from PET/CT. *Nucl. Med. Commun.* **2019**, *40*, 842–849. [[CrossRef](#)] [[PubMed](#)]
4. Zhu, L.; Yin, G.; Chen, W.; Li, X.; Yu, X.; Zhu, X.; Jiang, W.; Jia, C.; Chen, P.; Zhang, Y.; et al. Correlation between EGFR mutation status and F(18) -fluorodeoxyglucose positron emission tomography-computed tomography image features in lung adenocarcinoma. *Thorac. Cancer* **2019**, *10*, 659–664. [[CrossRef](#)]
5. Koyasu, S.; Nishio, M.; Isoda, H.; Nakamoto, Y.; Togashi, K. Usefulness of gradient tree boosting for predicting histological subtype and EGFR mutation status of non-small cell lung cancer on (18)F FDG-PET/CT. *Ann. Nucl. Med.* **2020**, *34*, 49–57. [[CrossRef](#)] [[PubMed](#)]
6. Garrido, P.; Conde, E.; de Castro, J.; Gómez-Román, J.J.; Felip, E.; Pijuan, L.; Isla, D.; Sanz, J.; Paz-Ares, L.; Lopez-Rios, F. Updated guidelines for predictive biomarker testing in advanced non-small-cell lung cancer: A National Consensus of the Spanish Society of Pathology and the Spanish Society of Medical Oncology. *Clin. Transl. Oncol.* **2020**, *22*, 989–1003. [[CrossRef](#)]
7. Doble, B.; John, T.; Thomas, D.; Fellowes, A.; Fox, S.; Lorgelly, P. Cost-effectiveness of precision medicine in the fourth-line treatment of metastatic lung adenocarcinoma: An early decision analytic model of multiplex targeted sequencing. *Lung Cancer* **2017**, *107*, 22–35. [[CrossRef](#)]
8. Johnston, K.M.; Sheffield, B.S.; Yip, S.; Lakzadeh, P.; Qian, C.; Nam, J. Comprehensive genomic profiling for non-small-cell lung cancer: Health and budget impact. *Curr. Oncol.* **2020**, *27*, 569–577. [[CrossRef](#)]
9. Yu, T.M.; Morrison, C.; Gold, E.J.; Tradonsky, A.; Arnold, R.J.G. Budget Impact of Next-Generation Sequencing for Molecular Assessment of Advanced Non-Small Cell Lung Cancer. *Value Health* **2018**, *21*, 1278–1285. [[CrossRef](#)]
10. Bewicke-Copley, F.; Arjun Kumar, E.; Palladino, G.; Korfi, K.; Wang, J. Applications and analysis of targeted genomic sequencing in cancer studies. *Comput. Struct. Biotechnol. J.* **2019**, *17*, 1348–1359. [[CrossRef](#)]
11. Schwartzberg, L.; Kim, E.S.; Liu, D.; Schrag, D. Precision Oncology: Who, How, What, When, and When Not? *Am. Soc. Clin. Oncol. Educ. Book* **2017**, *37*, 160–169. [[CrossRef](#)] [[PubMed](#)]
12. Hofman, V.; Rouquette, I.; Long-Mira, E.; Piton, N.; Chamorey, E.; Heeke, S.; Vignaud, J.M.; Yguel, C.; Mazieres, J.; Lepage, A.-L.; et al. Multicenter Evaluation of a Novel ROS1 Immunohistochemistry Assay (SP384) for Detection of ROS1 Rearrangements in a Large Cohort of Lung Adenocarcinoma Patients. *J. Thorac. Oncol.* **2019**, *14*, 1204–1212. [[CrossRef](#)] [[PubMed](#)]
13. D’Haene, N.; Le Mercier, M.; De Nève, N.; Blanchard, O.; Delaunoy, M.; El Housni, H.; Dessars, B.; Heimann, P.; Rummelink, M.; Demetter, P.; et al. Clinical Validation of Targeted Next Generation Sequencing for Colon and Lung Cancers. *PLoS ONE* **2015**, *10*, e0138245. [[CrossRef](#)] [[PubMed](#)]
14. Boellaard, R.; Delgado-Bolton, R.; Oyen, W.J.G.; Giammarile, F.; Tatsch, K.; Eschner, W.; Verzijlbergen, F.J.; Barrington, S.F.; Pike, L.C.; Weber, W.A.; et al. FDG PET/CT: EANM procedure guidelines for tumour imaging: Version 2.0. *Eur. J. Nucl. Med. Mol. Imaging.* **2015**, *42*, 328–354. [[CrossRef](#)]
15. Wanet, M.; Lee, J.A.; Weynand, B.; De Bast, M.; Poncelet, A.; Lacroix, V.; Coche, E.; Gregoire, V.; Geets, X. Gradient-based delineation of the primary GTV on FDG-PET in non-small cell lung cancer: A comparison with threshold-based approaches, CT and surgical specimens. *Radiother. Oncol.* **2011**, *98*, 117–125. [[CrossRef](#)]
16. 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](#)]
17. Hatt, M.; Majdoub, M.; Vallières, M.; Tixier, F.; Le Rest, C.C.; Groheux, D.; Hindle, E.; Martineau, A.; Pradler, O.; Hustinx, R.; et al. 18 F-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](#)]
18. 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](#)]
19. Image Features. Available online: https://ibsi.readthedocs.io/en/latest/03_Image_features.html (accessed on 26 September 2022).
20. Orhac, F.; Lecler, A.; Savatovski, J.; Goya-Outi, J.; Nioche, C.; Charbonneau, F.; Ayache, N.; Frouin, F.; Duron, L.; Buvat, I. How can we combat multicenter variability in MR radiomics? Validation of a correction procedure. *Eur. Radiol.* **2021**, *31*, 2272–2280. [[CrossRef](#)]
21. Black, W.C.; Armstrong, P. Communicating the significance of radiologic test results: The likelihood ratio. *AJR Am. J. Roentgenol.* **1986**, *147*, 1313–1318. [[CrossRef](#)]
22. Grimes, D.A.; Schulz, K.F. Refining clinical diagnosis with likelihood ratios. *Lancet* **2005**, *365*, 1500–1505. [[CrossRef](#)]
23. Calvayrac, O.; Pradines, A.; Pons, E.; Mazières, J.; Guibert, N. Molecular biomarkers for lung adenocarcinoma. *Eur. Respir. J.* **2017**, *49*, 601734. [[CrossRef](#)] [[PubMed](#)]
24. Toyooka, S.; Matsuo, K.; Shigematsu, H.; Kosaka, T.; Tokumo, M.; Yatabe, Y.; Ichihara, S.; Inukai, M.; Sueshisa, H.; Soh, J.; et al. The impact of sex and smoking status on the mutational spectrum of epidermal growth factor receptor gene in non small cell lung cancer. *Clin. Cancer Res.* **2007**, *13*, 5763–5768. [[CrossRef](#)] [[PubMed](#)]
25. Shigematsu, H.; Lin, L.; Takahashi, T.; Nomura, M.; Suzuki, M.; Wistuba, I.I.; Fong, K.M.; Lee, H.; Toyooka, S.; Shimizu, N.; et al. Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. *J. Natl. Cancer Inst.* **2005**, *97*, 339–346. [[CrossRef](#)]

26. Kosaka, T.; Yatabe, Y.; Endoh, H.; Kuwano, H.; Takahashi, T.; Mitsudomi, T. Mutations of the epidermal growth factor receptor gene in lung cancer: Biological and clinical implications. *Cancer Res.* **2004**, *64*, 8919–8923. [[CrossRef](#)]
27. Thomas, L.; Doyle, L.A.; Edelman, M.J. Lung cancer in women: Emerging differences in epidemiology, biology, and therapy. *Chest* **2005**, *128*, 370–381. [[CrossRef](#)]
28. Dogan, S.; Shen, R.; Ang, D.C.; Johnson, M.L.; D'Angelo, S.P.; Paik, P.K.; Brostowski, E.B.; Riely, G.J.; Kris, M.G.; Zakowski, M.F.; et al. Molecular epidemiology of EGFR and KRAS mutations in 3,026 lung adenocarcinomas: Higher susceptibility of women to smoking-related KRAS-mutant cancers. *Clin. Cancer Res.* **2012**, *18*, 6169–6177. [[CrossRef](#)]
29. Go, H.; Kim, D.W.; Kim, D.; Keam, B.; Kim, T.M.; Lee, S.H.; Heo, D.S.; Bang, Y.-J.; Chung, D.H. Clinicopathologic analysis of ROS1-rearranged non-small-cell lung cancer and proposal of a diagnostic algorithm. *J. Thorac. Oncol.* **2013**, *8*, 1445–1450. [[CrossRef](#)]
30. Warth, A.; Muley, T.; Dienemann, H.; Goeppert, B.; Stenzinger, A.; Schnabel, P.A.; Schirmacher, P.; Penzel, R.; Weichert, W. ROS1 expression and translocations in non-small-cell lung cancer: Clinicopathological analysis of 1478 cases. *Histopathology* **2014**, *65*, 187–194. [[CrossRef](#)]
31. Mitchell, K.G.; Parra, E.R.; Zhang, J.; Nelson, D.B.; Corsini, E.M.; Villalobos, P.; Moran, C.A.; Skoulidis, F.; Wistuba, I.I.; Fujimoto, J.; et al. LKB1/STK11 Expression in Lung Adenocarcinoma and Associations With Patterns of Recurrence. *Ann. Thorac. Surg.* **2020**, *110*, 1131–1138. [[CrossRef](#)]
32. Barta, J.A.; Powell, C.A.; Wisnivesky, J.P. Global Epidemiology of Lung Cancer. *Ann. Glob. Health* **2019**, *85*, 8. [[CrossRef](#)] [[PubMed](#)]
33. Chapman, A.M.; Sun, K.Y.; Ruestow, P.; Cowan, D.M.; Madl, A.K. Lung cancer mutation profile of EGFR, ALK, and KRAS: Meta-analysis and comparison of never and ever smokers. *Lung Cancer* **2016**, *102*, 122–134. [[CrossRef](#)] [[PubMed](#)]
34. Rivera, G.A.; Wakelee, H. Lung Cancer in Never Smokers. *Adv. Exp. Med. Biol.* **2016**, *893*, 43–57. [[PubMed](#)]
35. Lasnon, C.; Majdoub, M.; Lavigne, B.; Do, P.; Madelaine, J.; Visvikis, D.; Hatt, M.; Aide, N. (18)F-FDG PET/CT heterogeneity quantification through textural features in the era of harmonisation programs: A focus on lung cancer. *Eur. J. Nucl. Med. Mol. Imaging* **2016**, *43*, 2324–2335. [[CrossRef](#)] [[PubMed](#)]
36. Kaalep, A.; Burggraaff, C.N.; Pieplenbosch, S.; Verwer, E.E.; Sera, T.; Zijlstra, J.; Hoekstra, O.S.; Oprea-Lager, D.E.; Boellaard, R. Quantitative implications of the updated EARL 2019 PET-CT performance standards. *EJNMMI Phys.* **2019**, *6*, 28. [[CrossRef](#)] [[PubMed](#)]
37. Aide, N.; Lasnon, C.; Veit-Haibach, P.; Sera, T.; Sattler, B.; Boellaard, R. EANM/EARL harmonization strategies in PET quantification: From daily practice to multicentre oncological studies. *Eur. J. Nucl. Med. Mol. Imaging* **2017**, *44* (Suppl. S1), 17–31. [[CrossRef](#)]
38. Aide, N.; Salomon, T.; Blanc-Fournier, C.; Grellard, J.M.; Levy, C.; Lasnon, C. Implications of reconstruction protocol for histo-biological characterisation of breast cancers using FDG-PET radiomics. *EJNMMI Res.* **2018**, *8*, 114. [[CrossRef](#)]
39. Dai, W.; Luo, B.; Wu, Z.; Chen, J.; Feng, G.; Guan, P. A multi-center phase II study of nintedanib as second-line therapy for patients with advanced non-small-cell lung cancer in China. *Am. J. Cancer Res.* **2015**, *5*, 3270–3275.
40. Dearden, S.; Stevens, J.; Wu, Y.L.; Blowers, D. Mutation incidence and coincidence in non small-cell lung cancer: Meta-analyses by ethnicity and histology (mutMap). *Ann. Oncol.* **2013**, *24*, 2371–2376. [[CrossRef](#)]
41. Youlden, D.R.; Cramb, S.M.; Baade, P.D. The International Epidemiology of Lung Cancer: Geographical distribution and secular trends. *J. Thorac. Oncol.* **2008**, *3*, 819–831. [[CrossRef](#)]
42. Swanton, C.; McGranahan, N.; Starrett, G.J.; Harris, R.S. APOBEC Enzymes: Mutagenic Fuel for Cancer Evolution and Heterogeneity. *Cancer Discov.* **2015**, *5*, 704–712. [[CrossRef](#)] [[PubMed](#)]
43. Pelosi, G.; Pellegrinelli, A.; Fabbri, A.; Tamborini, E.; Perrone, F.; Settanni, G.; Busico, A.; Piccioni, B.; Testi, M.A.; Militti, L.; et al. Deciphering intra-tumor heterogeneity of lung adenocarcinoma confirms that dominant, branching, and private gene mutations occur within individual tumor nodules. *Virchows Arch.* **2016**, *468*, 651–662. [[CrossRef](#)] [[PubMed](#)]
44. Colomer, R.; Mondejar, R.; Romero-Laorden, N.; Alfranca, A.; Sanchez-Madrid, F.; Quintela-Fandino, M. When should we order a next generation sequencing test in a patient with cancer? *EClinicalMedicine* **2020**, *25*, 100487. [[CrossRef](#)] [[PubMed](#)]