Journal of Nuclear Medicine, published on November 15, 2019 as doi:10.2967/jnumed.119.228106

Pre-treatment ^{18}F-FDG PET/CT radiomics predict local recurrence in patients treated with stereotactic radiotherapy for early-stage non-small cell lung cancer: a multicentric study

Gurvan Dissaux[1,2], Dimitris Visvikis[2], Ronrick Da-ano[2], Olivier Pradier[1,2], Enrique Chajon[3], Isabelle Barillot[4], Loig Duverge[3], Ingrid Masson[5], Ronan Abgral[6], Maria-Joao Santiago Ribeiro[7], Anne Devillers[8], Amandine Pallardy [9], Vincent Fleury[10], Marc-André Mahé[5], Renaud De Crevoisier[3], Mathieu Hatt[2*], Ulrike Schick[1,2*]

1. Radiation Oncology department, University Hospital, Brest, France

2. LaTIM, INSERM, UMR 1101, Univ Brest, Brest, France

3. Radiotherapy Department, Centre Eugene Marquis, Rennes, France

4. Department of Radiation Oncology (CORAD), University Hospital, Tours, France

5. Department of Radiation Oncology, ICO, Saint-Herblain, France

6. Nuclear Medicine department, University Hospital, Brest, France

7. Nuclear Medicine department, University Hospital, Tours, France

8. Nuclear Medicine department, Centre Eugene Marquis, Rennes, France

9. Nuclear Medicine department, University Hospital, Nantes, France

10. Nuclear Medicine department, ICO, Saint-Herblain, France

*** both authors equally contributed**

Corresponding author: Gurvan DISSAUX

Radiation Oncology Department

CHRU Morvan, 2 avenue Foch

29609 Cedex, Brest, France

Tel: +33 7 86 71 81 05

E-mail: gurvan.dissaux@chu-brest.fr

Running title: PET/CT radiomics in SBRT for lung cancer

Wordcount: ~ 5090

ABSTRACT

Purpose: The aim of this retrospective multicentric study was to develop and evaluate a prognostic FDG PET/CT radiomics signature in early-stage non-small cell lung cancer (NSCLC) patients treated with stereotactic radiotherapy (SBRT). **Material and Methods**: Patients from 3 different centers (n=27, 29 and 8) were pooled to constitute the training set, whereas the patients from a fourth center (n=23) were used as the testing set. The primary endpoint was local control (LC). The primary tumour was semi-automatically delineated in the PET images using the Fuzzy locally adaptive Bayesian algorithm, and manually in the low-dose CT images. A total of 184 IBSI-compliant radiomic features were extracted. Seven clinical and treatment parameters were included. We used ComBat to harmonize radiomic features extracted from the four institutions relying on different PET/CT scanners. In the training set, variables found significant in the univariate analysis were fed into a multivariate regression model and models were built by combining independent prognostic factors. **Results**: Median follow-up was 21.1 (1.7 – 63.4) and 25.5 (7.7 – 57.8) months in training and testing sets respectively. In univariate analysis, none of the clinical variables, 2 PET and 2 CT features were significantly predictive of LC. The best predictive models in the training set were obtained by combining one feature from PET, namely information correlation 2 (IC2) and one from CT (Flatness), reaching a sensitivity of 100% and a specificity of 96%. Another model combining 2 PET features (IC2 and Strength), reached sensitivity of 100% and specificity of 88%, both with an undefined hazard ratio (HR) ($p<0.001$). The latter model obtained an accuracy of 0.91 (sensitivity 100%, specificity 81%), with a HR undefined ($p = 0.023$) in the testing set, however other models relying on CT radiomics features only or the combination of PET and CT features failed to validate in the testing set.

Conclusion: We showed that two radiomic features derived from FDG PET were independently associated with LC in patients with NSCLC undergoing SBRT and could be combined in an accurate predictive model. This model could provide local relapse-related information and could be helpful in clinical decision-making.

Keywords: PET/CT, radiomics, early stage NSCLC, stereotactic radiotherapy

INTRODUCTION

Non-small cell cancer (NSCLC) is usually associated with a poor prognosis. However, approximately 16% of patients present with early-stage cT1-T2 N0 disease at diagnosis (*1*). Over the past two decades, technological developments in target delineation, motion management, conformal treatment planning and daily image guidance have allowed the development of stereotactic body radiation therapy (SBRT) (*2*). SBRT uses stereotactic targeting to facilitate the accurate delivery of a short course of high-dose radiation to the target. SBRT has demonstrated high local control (LC) rates (85-90%), comparable to those obtained with surgery in multiple prospective trials (*3*) and is now a guideline-recommended treatment for patients with early stage NSCLC who are medically unfit or unwilling to undergo surgery (*4*). Amongst these patients, therapeutic results are nonetheless highly variable, and new predictive factors of response to SBRT are needed to better individualize treatment.

^{18}F-fluorodeoxyglucose (FDG) positron emission tomography/computed tomography (PET/CT) is the standard imaging tool for the initial staging and radiation treatment planning (*5*). PET/CT has also emerged as a prognostic tool in NSCLC, but parameters such as maximal standardized uptake value (SUV) and metabolic tumour volume have been inconsistently correlated with outcome (*6*).

Radiomic features are handcrafted metrics used to quantify tumour intensity, shape and heterogeneity, some of which have been shown to reflect intratumoural histopathological properties (*7*) and to predict patients' outcome in several pathologies including NSCLC when extracted from FDG PET, CT, or both (*8*).

We hypothesized that some radiomic features extracted from the FDG PET/CT images could have predictive value of recurrence in early-stage NSCLC patients treated with SBRT, and we aimed to evaluate this in a multicentric setting.

PATIENTS AND METHODS

Patients selection

Eighty-seven patients with NSCLC stage I-II and tumours diameters below 5 cm according to the 8th American Joint Committee on Cancer classification, treated with definitive curative SBRT from January 2012 to December 2016 at 4 French institutions (Rennes, Tours, Brest and Nantes) were retrospectively included.

All patients were required to have PET/CT imaging performed within 60 days of SBRT and at least 6 months of follow-up. Histological confirmation was not mandatory, but if a biopsy could not be performed due to contra-indication, a progression according to the Response Evaluation Criteria in Solid Tumour (RECIST) criteria on at least two serial CT imaging studies and/or an increased FDG uptake on PET/CT were necessary, according to recommendations (*9*).

Collected data included age and date of diagnosis, gender, performance status, histology when available, stage, tumour size as measured on CT according to RECIST criteria, status at last follow-up and PET/CT diagnostic images. Date and site of recurrence were also collected. Diagnosis of recurrences was based on CT findings (with confirmation of radiological progression on serial CT), and histological confirmation of relapse was not mandatory.

This study was approved by the Institutional Review Boards at each institution.

Treatment planning

The SBRT dose was prescribed according to each institution's protocol. Patients with peripherally located lesions received a median total radiation dose of 48 Gy (range 48 – 60) in 3 to 4 fractions. For central lesions, a median dose of 50 Gy (ranges 30-60 Gy) in 3 to 8 fractions was used (Supplemental table 1).

Follow-up

Clinical and radiological follow-up was performed at the treating institution. CTs were performed every 3 months in the first 2 years and thereafter every 6 months for another 3 years.

PET/CT image acquisition and image analysis

PET/CT images were acquired 60 +/- 5 min after FDG injection in accordance with the European Association of Nuclear Medicine guidelines (*10*). All PET images were corrected for attenuation using the acquired low-dose CT data. Acquisitions differed amongst the 4 institutions in terms of PET/CT scanner manufacturer and models, as well as in acquisitions protocols and reconstruction settings (Fig. 1 and Supplemental table 2).

Tumour delineation

The PET and the low-dose CT images were processed independently. All primary tumours were segmented by an expert radiation oncologist (GD), semi-automatically in the PET images using the fuzzy locally adaptive Bayesian (FLAB) algorithm, which has been shown to be robust with respect to differences in image acquisition and reconstruction settings (*11*), and manually in the low-dose CT images, with lung window setting (window level: -450 HU; window width: 1500 HU) using MiM Maestro® (MiM software Inc Cleveland, OH 44122) (Fig. 1).

Features extraction

Each tumour in both the PET and CT images was characterized with 92 radiomic features (shape, intensity and texture with fixed number of bins discretization into 64 bins and 3D merging strategy for matrices, see Supplemental table 2), compliant with the most up-to-date benchmark values of the image biomarkers standardization initiative (IBSI) (*12*).

Statistical analysis and modelling

For each patient, a total of 184 image features (92 in each modality, 7 clinical and histopathological parameters (age, gender, Gross-Tumor Volume (GTV) volume, stage, localization, OMS status, histology) and the biologically equivalent dose (BED) were included. In order to pool together radiomic features extracted from images acquired on the different PET/CT scanners and associated protocols, we used the *ComBat* harmonization method (*13*). After ComBat harmonization of radiomic features, patients from Rennes, Tours and Brest were

pooled to constitute the training set, whereas patients from Nantes were used as the testing set.

Primary endpoint was LC. Secondary endpoints were cancer specific survival (CSS), distant metastasis free survival (DMFS), recurrence free survival (RFS) and overall survival (OS). For the training set, all variables were tested using the univariate Cox proportional hazards model. Given the number of variables tested and the small number of events, a conservative significance level of $p<0.005$ was considered in order to reduce the risk of false discovery (*14*).

Receiver operating characteristic (ROC) curves were used to determine optimal cut-off values of significant variables using the Youden index. The resulting Kaplan-Meier curves for LC were compared using the log-rank test.

Because univariate selection ignores relationships between variables, Spearman rank correlation (ρ) was used to quantify correlations between the parameters with an Area Under the Curve (AUC) of the ROC curve above 0.7, in order to evaluate redundancy of potentially predictive variables (Supplemental Table 3). Cox regression models with the stepwise method were subsequently used for multivariate analysis by including only uncorrelated variables ($\rho<0.8$, a threshold arbitrarily chosen based on the usual categories suggested by Mukaka et al. (*15*), identified as significant in the univariate analysis, in order to identify independent factors that could be combined into multiparametric models. Such models were then built for binary classification of patients with both risk factors in one group versus the patients with none or only one of the risk factors.

The best models were evaluated in the testing set. Adjusted hazard ratios (HR) and the corresponding 95% confidence intervals (CI) were calculated.

In addition, correlations between the parameters identified in the multivariate analysis and standard metrics (*e.g.*, SUV_{max}, volume, etc.) were checked to avoid developing models that would end up being simply surrogates of usual variables (Supplemental Table 4).

All statistical analyses were performed using MedCalc Statistical Software version 18.5 (MedCalc Software, Ostend, Belgium; https://www.medcalc.org; 2018)

RESULTS

Patients and tumours characteristics

The flowchart (Supplemental Fig. 1) presents the patients selection. The 87 patients were split into a training set which included 64 patients from Brest (n=8), Rennes (n=27) and Tours (n=29), whereas 23 patients from Nantes constituted the testing set. Patients, tumours and treatment characteristics are provided in Table 1 and supplemental table 5.

All clinical parameters were not statistically significant between the training and testing sets. In 51 patients (59 %), the diagnosis of NSCLC was confirmed histologically (34 in the training cohort and 17 patients in the testing cohort).

In the training set, the median SBRT dose delivered was 54 Gy in 4 fractions (range 30– 60 Gy in 3 to 8 fractions, median BED = 150). In the testing set, the median SBRT dose delivered was 48 Gy in 4 fractions (median BED = 105.6 Gy).

There was no significant difference between the median SUV_{max} in training and testing sets, or in patients with biopsy proven NSCLC (median 7.5, range 2.6-36.2) compared to those without histology (median 7.7, range 2.5-18.23).

Outcome

TRAINING SET. In the training set, median follow-up was 21.1 months (range 1.7 – 63.4 months). Progression or disease recurrence occurred in 15 patients (24%) after a median follow-up of 37 (range 22 - 41) months. Eleven patients had distant recurrence (17%), none had regional recurrence and four (6%) had a local failure. The 2-year LC, OS, CSS, DMFS, and RFS rates were 90%, 75%, 89%, 75%, and 69%, respectively.

TESTING SET. In the testing set, median follow-up was 25.5 months (range 7.7 – 57.8 months). Progression or recurrence occurred in 9 patients (39%), seven (23%) had a distant recurrence, four (17%) had regional recurrence and two (9%) had a local failure. The 2-year LC, OS, CSS, DMFS, and RFS rates were 87%, 100%, 100%, 75%, and 72%, respectively.

Development of the radiomic signature in the training set

LOCAL CONTROL. In univariate analysis, no clinical parameter was statistically correlated with LC. BED10 (>151.2 Gy) showed association with worse LC without however reaching statistical significance with a HR of 10 (1 – 95), p = 0.016) (table 2).

Three CT-derived features and two PET-derived features reached AUC > 0.7 (Supplementary table 4). In univariate analysis, Flatness and Shade from CT were associated with LC (p<0.001, HR undefined, and p=0.003 HR of 13 (1.1-168), respectively). Information correlation 2 (IC2) from Gray-level Co-occurrence Matrix (GLCM) and texture strength (Strength) from Neighborhood Gray-Tone Difference Matrix from PET remained significant, with HR undefined for both (p=0.005 and p=0.001, respectively).

The model combining these two PET features (with cut-off values of 0.89 and 45.11 for IC2 and Strength respectively) reached an accuracy of 0.94 (sensitivity 100%, specificity 88%) to predict LC, with a HR undefined (p<0.001) (Fig. 2A). The estimated 3-year LC rates between patients with low versus high values for this model were 100% and 62.5% respectively (Fig. 2B). A PET/CT model combining $IC2_{PET}$ and $Flatness_{CT}$ (with cut-off values of 0.89 and 0.73 respectively) reached an accuracy of 0.98 (sensitivity 100%, specificity 96%) to predict local control with a HR undefined (p<0.001) (Fig. 3). The estimated 3-year LC rates between patients with low versus high PET/CT model were 100% and 60% respectively.

In the multivariate analysis, PET ($IC2_{PET}$ and $Strength_{PET}$) and PET/CT ($IC2_{PET}$ and $Flatness_{CT}$) models remained statistically significant with HR of 31 (3 – 369), p=0.007 and 38 (3 – 449), p=0.004, respectively (table 2).

OVERALL AND CANCER-SPECIFIC SURVIVAL. In univariate analysis, none of the investigated parameters (clinical or radiomic) were associated with OS or CSS. BED_{10}, with a 150 Gy cut-off, showed a trend for CSS (HR of 6 (1 – 49), p=0.07).

DISTANT METASTASES FREE SURVIVAL AND RECURRENCE FREE SURVIVAL. In univariate analysis, BED_{10} (>150 Gy) remained statistically correlated with DMFS and RFS with a HR of 8 (2-

27), p<0.001 and 6 (2-18), p<0.001, respectively (table 2). None of the radiomic features was associated with DMFS or RFS.

Evaluation of the radiomic signatures in the testing set

The PET-only signature combining $IC2_{PET}$ and $Strength_{PET}$ reached an accuracy of 0.91 for LC (sensitivity 100%, specificity 81%) (Figs. 4A-B), with a HR undefined (p = 0.023). The PET/CT signature combining $IC2_{PET}$ and $Flatness_{CT}$ failed to reach statistical significance.

The predictive power of SBRT regarding DFFS and RFS was not confirmed.

DISCUSSION

Our results are in line with previous studies in NSCLC that exploited FDG PET/CT to predict outcome, showing that tumour uptake distribution quantitatively characterized by FDG-PET radiomics is associated with response to external beam radiotherapy (*16*).

In this study, we identified a PET radiomic signature combining two textural features, namely IC2 and Strength that has predictive power regarding the efficacy of SBRT in the treatment of early-stage NSCLC. We showed that the higher these features were, the worse local control was, reflecting tumours heterogeneity on staging PET/CT of patients with lung cancer eligible to SBRT.

Radiomic features, exhibit variable sensitivity to the acquisition and reconstruction parameters (*17*). These can vary considerably from one institution to another or even within an institution relying on different scanners (e.g., Nantes in the present study), and therefore, validating radiomics-based models in a multicentric context is often challenging. This is one of the identified limitations for a broader transfer of radiomics to clinical practice. We recently confirmed that the ComBat method could successfully harmonize radiomics features extracted from PET and MR images obtained with different acquisition and reconstruction parameters in order to facilitate validation of radiomics signatures in a multicentric setting (*18*). In our present work, we further confirmed the interest of ComBat (Supplemental Fig. 2). Indeed, without harmonization the radiomic features had lower predictive power (Supplemental table 6).

Studies focusing on early stage lung tumours treated with SBRT are sparse. A first retrospective study in a cohort of 101 patients treated with SBRT showed that peak standardized uptake value and radiomic features could predict distant recurrence with a C-index of 0.71, higher than SUV_{max} or tumour volume alone (*19*). This work considered only PET images and was monocentric, although two different scanners were used, which was not accounted for in the radiomic features computation or statistical analysis. Another study reported $entropy_{GLCM}$ as independently associated with LC in a series of 45 patients treated with SBRT (*20*). Oikonomou *et al.* performed a larger analysis of 150 patients combining CT and PET parameters, but did not specifically report on LC. They used a manual PET segmentation which suffers from high inter- and intra-observer variability (*21*). Furthermore, they only tested 21 texture parameters extracted on 2 dimensions. Finally, Lovinfosse *et al*. identified $dissimilarity_{GLCM}$ from FDG PET in 63 patients as associated with DSS and DFS but not OS (*22*). All of these studies were retrospective, monocentric, and none followed a training/testing scheme, which can be crucial as we showed here that features predictive in the training set may not be validated in the testing set (the CT features in our case). In addition, most of them investigated a comparatively small number of non IBSI-compliant radiomic features and relied on different segmentation and/or intensity discretization schemes, which may contribute in explaining why different features were identified amongst these studies. Reproducibility and comparison between radiomics studies outside of the IBSI standardization framework is impossible (*23*).

In most radiomics studies, the number of variables is often greater than the number of patients, potentially leading to a high risk of false discovery (*14*). In order to address this, we relied on a more conservative *p*-value threshold in the univariate analysis of the training cohort (0.005 instead of the usual 0.05), and we evaluated the trained models in an separate dataset. With this approach, we found a LC signature consisting of 2 PET features that yielded a high prognostic performance for LC. The retained features showed low ($|\rho|<0.01$ for CT Flatness) to moderate ($|\rho|<0.64$ for PET Strength) rank correlation with the associated tumor volume (supplemental table 3).

Overall, 6 patients out of 87 (7%) developed a local relapse later on. This low rate of infield relapse is in accordance with the literature (*24*).

Our study has several limitations. First, histological confirmation was not available for 43% of the patients, leaving a doubt that some of them may have been treated for a benign disease. However, findings from surgical studies have shown that likelihood of benign diagnosis in patients having a new or growing lesion on CT with a corona radiate sign and local ^{18}F-FDG uptake is less than 4% (*25*). In addition, data from the literature show that therapeutic results in patients without histology are similar to those with histology or those that benefit from surgical resection (*26*). In our present work, we showed no difference between patients with or without histology (Supplemental Table 7). Another limitation is that all PET images were acquired in free breathing condition, which could influence radiomics although some features can be robust when compared between free breathing and with optimal respiratory gating (*27*). The segmentation on the CT images was performed manually by a single expert in the same way as he routinely delineates the GTV to plan stereotactic treatment, which prevents an evaluation of the inter-observer reliability of the CT features (*28*). Of note, the signature including a CT feature could not be validated in the testing set, contrary to the PET-only signature that combines features extracted from semi-automatically delineated tumour volumes with the FLAB algorithm, which allows reducing inter-observer variability. Several alternative segmentation methods can provide similar performance (*29*). It should also be emphasized that the CT component of PET/CT exploited in the present study is not the same resolution and quality as a planning or a diagnostic CT, and was also without contrast-enhancement. This can also contribute in explaining the lower predictive value of CT features compared to PET. Finally, our study was retrospective (as most radiomics studies) and included a limited number of patients, which is also a limitation to apply ComBat despite its proven robustness for small samples, since the number of patients per batch was small. The small number of events may also have limited the statistical significance of the results, and the representativeness of the model assessment (e.g., some of the HR in our results could not be determined due to 100% sensitivity). A post-hoc power analysis regarding an AUC of 0.9 compared to the null hypothesis (AUC 0.5) with alpha risk of 0.05 and power of 0.8 shows the training cohort is sufficient (4 events for 60 patients as

required) but the testing cohort is underpowered (2 events for 21 patients instead of 4 events for 42 patients). Finally, some patients had short follow-up (5 patients had follow-up less than 1 year), and could thus still develop recurrence despite being predicted as low risk, which would decrease the overall performance of our model.

CONCLUSION

In this multicenter retrospective study, we showed that imaging features derived from PET/CT were independent predictive factors of LC in patients with NSCLC undergoing SBRT. These features could provide recurrence-related information and could be helpful in clinical decision-making, especially regarding dose escalation. Our findings need to be confirmed in a larger cohort, which is currently being collected at our institution and other centers.

Data Availability Statement

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Acknowledgements

Funding

This work is part of the MuMoFraT project supported by the Canceropôle du Grand Ouest and the regions of Bretagne, Pays de la Loire and Centre - Val de Loire.

The PhD of R. Da-ano is funded by the H2020 Marie Curie ETN PREDICT project.

COMPLIANCE WITH ETHICAL STANDARDS

Conflict of interest

No potential conflicts of interest relevant to this article exist.

Ethical approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent

Informed consent was obtained from all individual participants included in the study.

KEY POINT

Question: Can pre-treatment 18-FDG PET/CT radiomics predict local recurrence in patients treated with stereotactic radiotherapy for early-stage non-small cell lung cancer?

Pertinent findings: We showed the ComBat harmonization method allowed to efficiently pool radiomic features extracted from the 4 clinical centers in order to train and validate a PET radiomics model combining 2 image features that achieves high accuracy in predicting local recurrence, especially in comparison with clinical factors that were not predictive.

Implications for patient care: a simple PET derived radiomics signature (2 features) may provide recurrence-related information and could be helpful in clinical decision-making, especially regarding dose escalation.

REFERENCES

1. Howlader N NA, Krapcho M, Garshell J, Miller D, Altekruse SF, Kosary CL, Yu M, Ruhl J, Tatalovich Z,Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA (eds). SEER Cancer Statistics Review, 1975-2012, National Cancer Institute. Bethesda, MD, https://seer.cancer.gov/csr/1975_2012/, based on November 2014 SEER data submission, posted to the SEER web site, April 2015. Accessed on 31st August 2019

2. Timmerman R, Paulus R, Galvin J, et al. Stereotactic body radiation therapy for inoperable early stage lung cancer. *JAMA.* 2010;303:1070-1076.

3. Chang JY, Senan S, Paul MA, et al. Stereotactic ablative radiotherapy versus lobectomy for operable stage I non-small-cell lung cancer: a pooled analysis of two randomised trials. *Lancet Oncol.* 2015;16:630-637.

4. Schneider BJ, Daly ME, Kennedy EB, et al. Stereotactic body radiotherapy for early-stage non-small-cell lung cancer: American society of clinical oncology endorsement of the American society for radiation oncology evidence-based guideline. *J Clin Oncol.* 2018;36:710-719.

5. Ettinger DS, Wood DE, Aisner DL, et al. Non-small cell lung cancer, version 5.2017, NCCN clinical practice guidelines in oncology. *J Natl Compr Canc Netw.* 2017;15:504-535.

6. Sharma A, Mohan A, Bhalla AS, et al. Role of various metabolic parameters derived from baseline 18F-FDG PET/CT as prognostic markers in non-small cell lung cancer patients undergoing platinum-based chemotherapy. *Clin Nucl Med.* 2018;43:e8-e17.

7. Rios Velazquez E, Parmar C, Liu Y, et al. Somatic mutations drive distinct imaging phenotypes in lung cancer. *Cancer Res.* 2017;77:3922-3930.

8. Desseroit MC, Visvikis D, Tixier F, et al. Development of a nomogram combining clinical staging with (18)F-FDG PET/CT image features in non-small-cell lung cancer stage I-III. *Eur J Nucl Med Mol Imaging.* 2016;43:1477-1485.

9. Videtic GMM, Donington J, Giuliani M, et al. Stereotactic body radiation therapy for early-stage non-small cell lung cancer: Executive summary of an ASTRO Evidence-Based Guideline. *Pract Radiat Oncol.* 2017;7:295-301.

10. Boellaard R, Delgado-Bolton R, Oyen WJ, et al. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. *Eur J Nucl Med Mol Imaging.* 2015;42:328-354.

11. Hatt M, Cheze Le Rest C, Albarghach N, Pradier O, Visvikis D. PET functional volume delineation: a robustness and repeatability study. *Eur J Nucl Med Mol Imaging.* 2011;38:663-672.

12. Zwanenburg A VM, Löck S. Image biomarker standardisation initiative - feature definitions. . 2017. https://arxiv.org/abs/1612.07003. Accessed on 31st August 2019

13. 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.

14. Chalkidou A, O'Doherty MJ, Marsden PK. False discovery rates in PET and CT studies with texture features: A systematic review. *PLoS One.* 2015;10:e0124165.

15. Mukaka MM. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Med J.* 2012;24:69-71.

16. Desseroit MC, Tixier F, Weber WA, et al. Reliability of PET/CT shape and heterogeneity features in functional and morphologic components of non-small cell lung cancer tumors: A repeatability analysis in a prospective multicenter cohort. *J Nucl Med.* 2017;58:406-411.

17. Galavis PE, 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.

18. Lucia F, Visvikis D, Vallieres M, et al. External validation of a combined PET and MRI radiomics model for prediction of recurrence in cervical cancer patients treated with chemoradiotherapy. *Eur J Nucl Med Mol Imaging.* 2018.

19. Wu J, Aguilera T, Shultz D, et al. Early-stage non-small cell lung cancer: quantitative imaging characteristics of (18)F fluorodeoxyglucose PET/CT allow prediction of distant metastasis. *Radiology.* 2016;281:270-278.

20. Pyka T, Bundschuh RA, Andratschke N, et al. Textural features in pre-treatment [F18]-FDG-PET/CT are correlated with risk of local recurrence and disease-specific survival in early stage NSCLC patients receiving primary stereotactic radiation therapy. *Radiat Oncol.* 2015;10:100.

21. Oikonomou A, Khalvati F, Tyrrell PN, et al. Radiomics analysis at PET/CT contributes to prognosis of recurrence and survival in lung cancer treated with stereotactic body radiotherapy. *Sci Rep.* 2018;8:4003.

22. Lovinfosse P, Janvary ZL, Coucke P, et al. FDG PET/CT texture analysis for predicting the outcome of lung cancer treated by stereotactic body radiation therapy. *Eur J Nucl Med Mol Imaging.* 2016;43:1453-1460.

23. Vallieres M, Zwanenburg A, Badic B, Cheze Le Rest C, Visvikis D, Hatt M. Responsible radiomics research for faster clinical translation. *J Nucl Med.* 2018;59:189-193.

24. Murray L, Ramasamy S, Lilley J, et al. Stereotactic ablative radiotherapy (SABR) in patients with medically inoperable peripheral early stage lung cancer: Outcomes for the first UK SABRcohort. *Clin Oncol (R Coll Radiol).* 2016;28:4-12.

25. Girard N, Mornex F. [Stereotactic radiotherapy for non-small cell lung cancer: From concept to clinical reality. 2011 update]. *Cancer Radiother.* 2011;15:522-526.

26. Haasbeek CJ, Lagerwaard FJ, Slotman BJ, Senan S. Outcomes of stereotactic ablative radiotherapy for centrally located early-stage lung cancer. *J Thorac Oncol.* 2011;6:2036-2043.

27. Grootjans W, Tixier F, van der Vos CS, et al. The impact of optimal respiratory gating and image noise on evaluation of intratumor heterogeneity on 18F-FDG PET imaging of lung cancer. *J Nucl Med.* 2016;57:1692-1698.

28. Eminowicz G, McCormack M. Variability of clinical target volume delineation for definitive radiotherapy in cervix cancer. *Radiother Oncol.* 2015;117:542-547.

29. Hatt M, Laurent B, Ouahabi A, et al. The first MICCAI challenge on PET tumor segmentation. *Med Image Anal.* 2018;44:177-195.

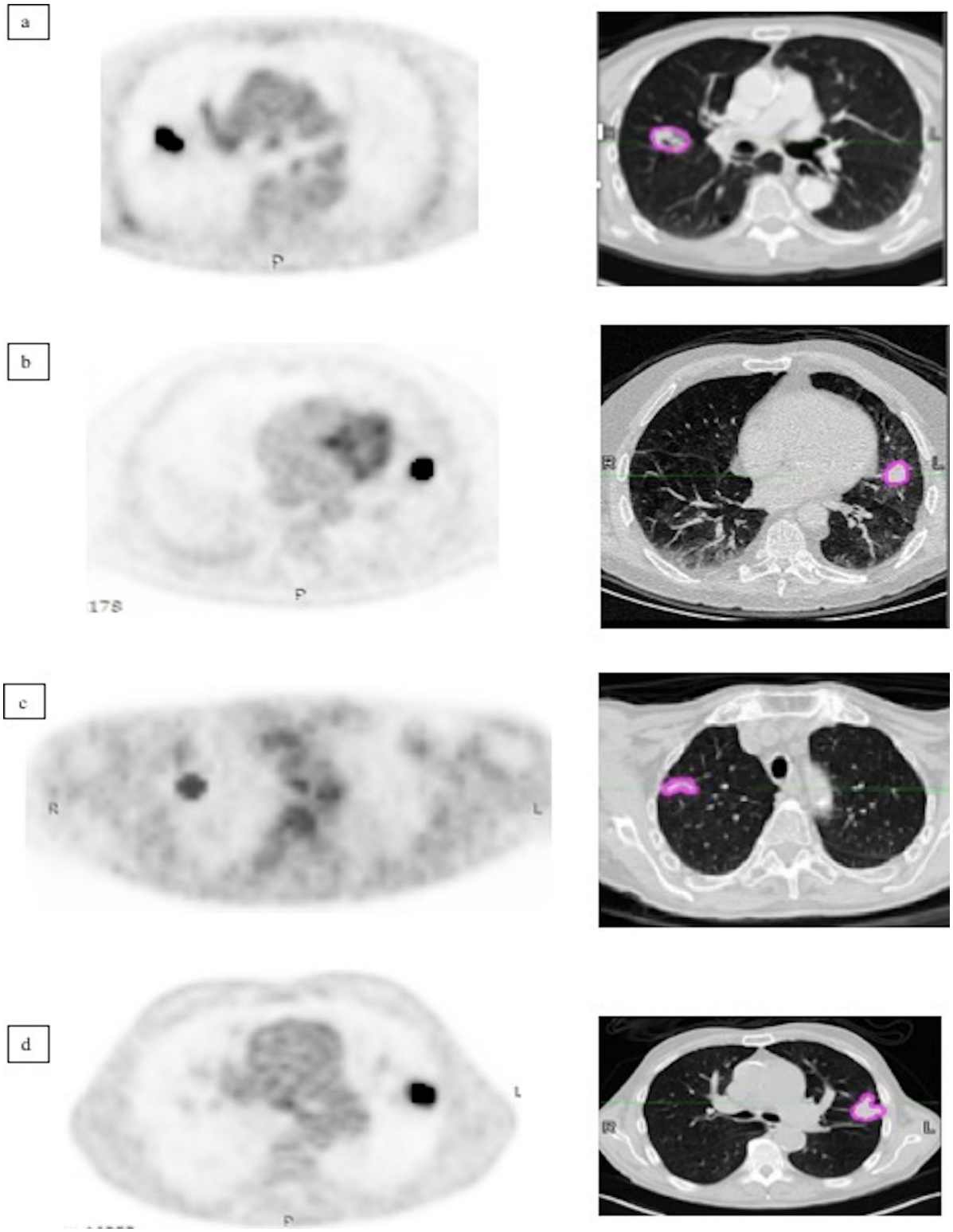
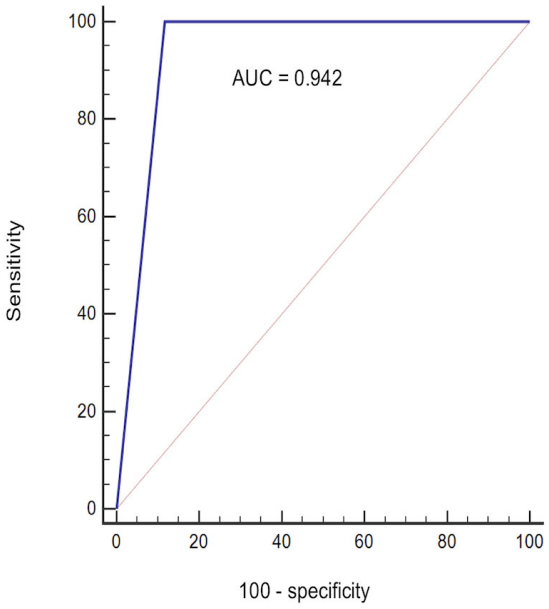


Fig. 1: Volume Of Interest on each PET/CT model: a. Brest (Biograph), b. Rennes (Discovery ST), c. Tours (Ingenuity), d. Nantes (Biograph) with semi-automatically fuzzy locally adaptive Bayesian (FLAB) segmentation on corrected PET and corresponding manual gross-tumor volume (GTV) segmentation on CT with lung window setting



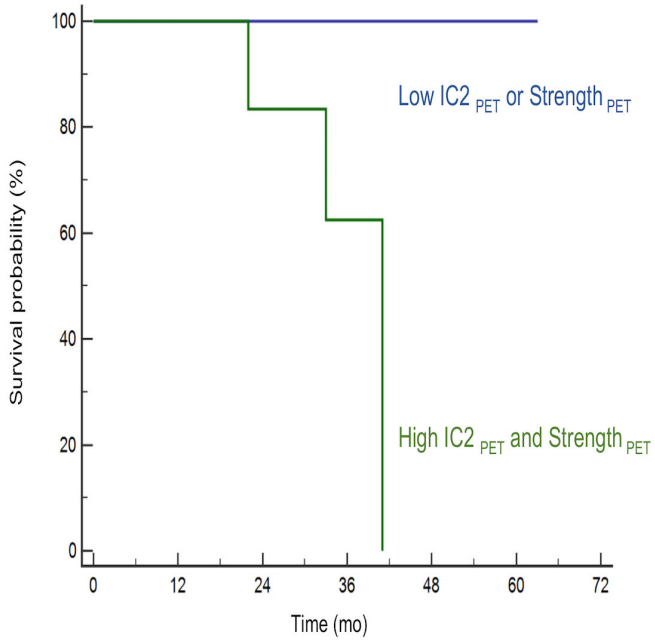


Fig. 2A: Receiver Operating Curve analysis for the prediction of local control according to $IC2_{PET}$ and $Strength_{PET}$ in the training cohort

Fig. 2B: Kaplan-Meier curve of local control based on radiomics signature $IC2_{PET}$ and $Strength_{PET}$ in the training cohort

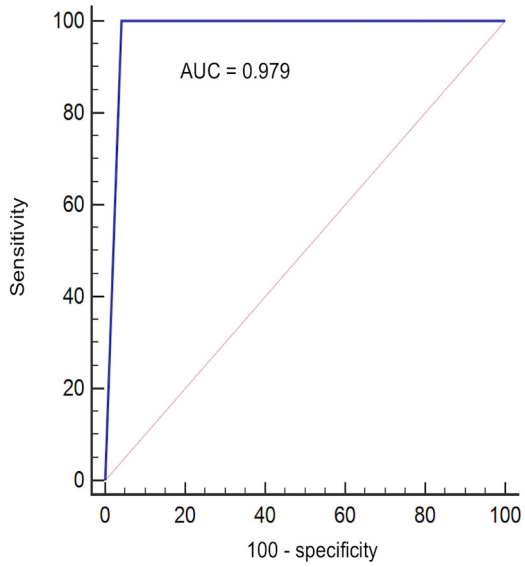
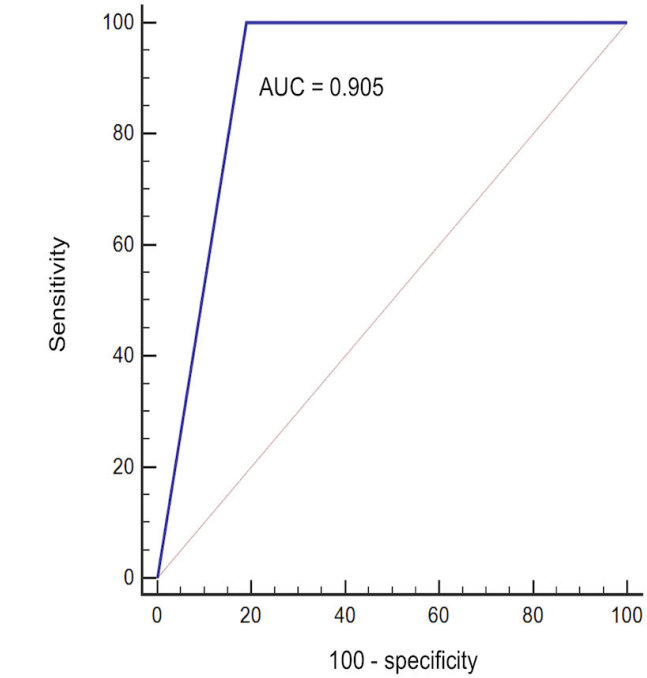


Fig. 3: Receiver Operating Curve analysis for the prediction of local control according to the model combining $IC2_{PET}$ and $Flatness_{CT}$ in the training cohort



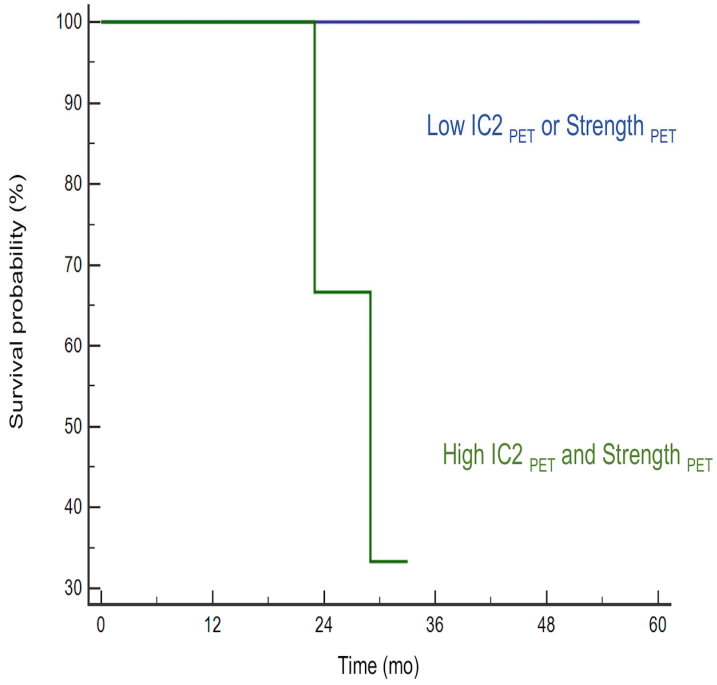


Fig. 4A: Receiver Operating Curve analysis for the prediction of local control according to the model combining IC2 PET and Strength PET in the testing cohort
Fig. 4B: Kaplan-Meier curve of local control based on the radiomics model combining IC2 PET and Strength PET in the testing cohort

TABLES

Table1: Patients and tumours characteristics

Variables	Training set (n=64)		Testing set (n=23)		*p*-value
Age (year) (median, range)	72	45 - 87	69	52 - 85	0.589
Tumor volume (cm^3)	133.3	8.3 – 946.9	149.7	5.7 – 987.2	0.946
T stage					
T1	13	20.3	4	17.4	0.864
T2	51	79.7	19	82.6	
Gender					
Male	48	75	19	82.6	0.821
Female	16	25	4	17.4	
Histology					
Adenocarcinoma	25	39	12	52.2	0.742
Other histologies	9	14	5	21.7	
Unknown	30	47	6	26.1	
WHO performance status					
0-1	49	76.6	22	95.6	0.435
>1	15	23.4	1	4.4	
Tumor location					
Peripheral	55	85.9	21	91.3	0.558
Central	9	14.1	2	8.7	
BED10					
<150 Gy	41	64	17	73.9	0.599
>150 Gy	23	36	6	26.1	
SUV_{max}	7.76	2.5 – 36.2	7.79	2.5 – 18.6	0.538

Abbreviations: WHO= World Health Organization, BED= biologically equivalent dose, SUV=Standard Uptake Value

Table 2: Uni- and multivariate analysis for local control in the training set

	Univariate analysis			Multivariate analysis		
Variables	HR	95% CI	p value	HR	95% CI	p value
Clinical variable						
Age (year) (<72 y *vs.* ≥ 72 y)	-	-	0.975			
Tumor volume (>132 cm^3 *vs*. ≤132 cm^3)	-	-	0.059			
TNM (T2 *vs*. T1)	6	0.1 – 405	0.078			
Gender (male *vs.* female)	3	0.3 – 29	0.251			
Histopathologic type (others *vs*. ADK)	3	0.4 – 21	0.297			
WHO performance status (≤1 *vs.* 2)	-	-	0.420			
Localisation (peripheral *vs*. central)	-	-	0.793			
BED10 (>151.2 Gy *vs.* ≤151.2 Gy)	-	-	0.067			
Radiomic signatures						
PET signature (IC2 and Strength)	-	-	**<0.001**	30.75	2.56 - 369.11	**0.007**
PET/CT signature (PET IC2 and CT Flatness)	-	-	**<0.001**	37.5	3.13 – 448.61	**0.004**

Abbreviations: ADK=adenocarcinoma, WHO= World Health Organization, BED=biologically equivalent dose.