Training and validation of a robust PET radiomic-based index to predict distant-relapse-free-survival after radio-chemotherapy for locally advanced pancreatic cancer

Martina Mori[1,2], Paolo Passoni[3], Elena Incerti[4], Valentino Bettinardi[4], Sara Broggi[1], Michele Reni[5], Phil Whybra[6], Emiliano Spezi[6,7], Elena G Vanoli[4], Luigi Gianolli[4], Maria Picchio[2,4], Nadia G Di Muzio[2,3], Claudio Fiorino[1]

1) Medical Physics, San Raffaele Scientific Institute, Milano, Italy.
2) Vita-Salute San Raffaele University, Milan, Italy.
3) Radiotherapy, San Raffaele Scientific Institute, Milano, Italy.
4) Nuclear Medicine, San Raffaele Scientific Institute, Milano, Italy.
5) Oncology, San Raffaele Scientific Institute, Milano, Italy.
6) School of Engineering, Cardiff University, Cardiff UK
7) Department of Medical Physics, Velindre Cancer Centre, Cardiff, UK

ABSTRACT

Purpose: To assess the value of 18F-Fluorodeoxyglucose (18F-FDG) PET Radiomic Features (RF) in predicting Distant Relapse Free Survival (DRFS) in patients with Locally Advanced Pancreatic Cancer (LAPC) treated with radio-chemotherapy.

Materials & Methods: One-hundred-ninety-eight RF were extracted using an IBSI (Image Biomarker Standardization Initiative) consistent software from pre-radiotherapy images of 176 LAPC patients, treated with moderate hypo-fractionation (44.25Gy, 2.95Gy/fr). Tumors were segmented by applying a previously validated semi-automatic method. One-hundred-twenty-six RF were excluded due to poor reproducibility and/or repeatability and/or inter-scanner variability. The original cohort was randomly split in a training (n=116) and a validation (n=60) group. Multi-variable Cox regression was applied on the training group, including only independent RF in the model. The resulting radiomic index was tested in the validation cohort. The impact of selected clinical variables was also investigated.

Results: The resulting Cox model included two RF of first order: Center of Mass Shift (COMshift) and 10th Intensity percentile (P_{10}) (p=0.0005, HR=2.72, 95%CI=1.54-4.80), showing worse outcome for patients with lower COMshift and higher P_{10}. Once stratified by quartile values (< lowest quartile *vs* > highest quartile *vs* the remaining), the index properly stratified patients according to their DRFS (p=0.0024, log-rank test). Performances were confirmed in the validation cohort (p=0.03, HR=2.53, 95%CI=0.96-6.65). The addition of clinical factors did not significantly improve models' performances.

Conclusions: A radiomic-based index including only two robust PET-RF predicted DRFS of LAPC patients after radio-chemotherapy. Current results could find relevant applications in treatment personalization of LAPC. A multi-institution independent validation has currently been planned.

Introduction

Pancreatic Adenocarcinoma is associated with high mortality and shows increasing incidence in most countries [1,2]. In recent years, advancements concerning cancer-specific outcome have been minimal, with 5-year overall survival (OS) rates below 10% [1]. Surgery, as a treatment of choice, leads to 5-year OS rates up to 20-25%; nevertheless, a large fraction of patients is unfit to surgery at diagnosis due to the stage of the disease or other concomitant exclusion criteria [3]. In fact, the majority of patients are diagnosed in a locally advanced (LAPC) or metastatic stage, for which 5-year survival is poor, while patients diagnosed with local disease, have better prognosis. [4]. Chemo-radiotherapy (CRT) is a frequently adopted solution but with the drawback of some risk of severe toxicity and still poor survival, with median OS around 5-15 months, 2-year OS below 30% and prevalent pattern of distant relapse [5,6]. Due to this, the availability of reliable models to predict Distant Relapse Free Survival (DRFS) could have a positive impact on treatment personalization of LAPC: for example, OS could be improved with more intensive local therapies for patients at lower risk of early distant relapse; on the other hand, radiotherapy could be avoided (or doses reduced aiming to achieve just a palliative effect) in patients candidate to rapidly develop metastatic disease. Unfortunately, there is an evident lack of reliable models to predict DRFS after CRT. At present, predictors of survival in patients with pancreatic cancer include tumor size, level of differentiation and lymph node status [9]. Molecular biomarkers for pancreatic cancer can be categorized into prognostic and predictive markers, such as the well-known Ca19.9 (carbohydrate antigen 19.9) [7]: CA 19.9 is the only Food and Drug Administration–approved biomarker for LAPC, with demonstrated clinical usefulness for therapeutic monitoring and early detection of recurrent disease after treatment [8]. However, it has important limitations, not being a specific biomarker for pancreatic cancer since its level may be elevated in other conditions and some patients do not synthetize it [9]. In this context, functional imaging techniques have been explored to assess non-invasive biomarkers; metabolic imaging, such as PET with 18F-fluorodeoxyglucose (18F-FDG-PET), is particularly promising. Basic PET features provided additional information to predict pancreatic tumor behavior: examples include maximum and mean standardized uptake value (SUVmax, SUV mean), metabolic tumour volume (MTV), and total lesion glycolysis (TLG) [10-12].

More advanced quantitative image-based analyses dealt with the radiomic field: radiomics aims to convert medical images into quantitative descriptors of oncologic tissues providing extensive information about intensity, shape, size/volume and texture of tumors [13]. PET images have been garnering more interest because the intensity heterogeneity is expected to be directly associated to tumor biology [14,15] and few studies tried to explore the pancreatic cancer scenario [16-18]. On the other hand, reliability and repeatability of radiomic features (RF) in the pancreatic cancer context is a crucial issue, as many RF are expected to be highly unreliable, mostly due to the expected uncertainty in tumor delineation, and this needs specific investigations [19]. Based on the available data of a large Institutional cohort of LAPC patients treated with CRT, selected robust RF were tested as potentially predictive of DRFS. In short, the aims of current investigation were:

a) Assessing a robust PET radiomic index to predict DRFS in a training cohort of patients treated with the same radiotherapy protocol with CRT, putting attention to the generalizability of the resulting index.
b) Validating the index on a validation cohort treated in the same Institute with the same protocol.

Materials and methods:

Patients' population
The current retrospective study enrolled 176 patients with histologically confirmed LAPC who underwent 18F-FDG PET/Computed Tomography (CT) before radiotherapy, from 2006 to 2018 at San Raffaele Institute in Milan. All patients were excluded from surgery because were judged unresectable according to NCCN Guidelines (version 2.2018) [20]. Criteria for inclusion were as follows: (a) histological diagnosis of pancreatic adenocarcinoma, (b) locally advanced disease, (c) bidimensional measurable disease, (d) Karnofsky performance status scale >50 and (e) age >18 years. Most details of radiotherapy and chemotherapy delivery at our Institute were reported elsewhere [21,22]. In short, all patients with stage III or patients with stage IV still in clinical complete response four months after the end of induction chemotherapy were treated with chemo-radiation (CRT), according to our institutional policy [23]. The target volume consisted of tumor and radiologically involved/PET positive lymph-nodes [21]. Induction chemotherapy schedules changed over time: for most patients it consisted of four-six cycles of four-drug combinations: cisplatin, epirubicin, 5-fluorouracil or capecitabine, gemcitabine (acronyms: PEFG of PEXG). All patients were treated with IMRT in a moderate hypo-fractionation approach, delivering 44.25Gy in 15 fractions with Helical Tomotherapy (Accuray Inc) or VMAT (Varian Linac DHX, 6MV X-Rays). For selected patients, concomitant boost up to 48Gy (median dose) to the portion of tumor infiltrating the peri-pancreatic vessels was delivered [22]. Concomitant chemotherapy consisted of capecitabine, 1250 mg/m^2/day. PET/CT was always performed after the completion of induction chemotherapy. The timing and the site of distant progression and death were prospectively registered; for stage IV patients, the occurrence of disease in a different metastatic site was defined as distant relapse. The study was approved by IRB and all patients provided written informed consent before undergoing 18F-FDG PET/CT. Main patients' characteristics are reported in Table 1.

Image Acquisition, settings parameters
Three different scanners were used for image acquisition (Discovery-ST, Discovery-STE, and Discovery-690, General Electric Medical Systems, Milwaukee, WI, USA). According to a clinical oncological protocol fasting the patients for more than 6h, 18F-FDG dose was 370 MBq, and static emission images were acquired on average 60 minutes after the tracer injection. Attenuation correction and image fusion were performed using CT data, and PET data were iteratively reconstructed with attenuation correction and axial, sagittal and coronal slices generated. Details are reported elsewhere [10].

Image sampling
All images were resampled at cubic voxels of 3×3×3 mm^3 with a bilinear interpolation using an automatic workflow on purpose developed in a dedicated commercial software (MIM Software Inc v.6.5.5). This procedure was implemented to reduce directional bias when the voxel sizes are not already isotropic [24], according to the specific recommendation of the International Biomarker Standardization Initiative (IBSI), in order to allow comparison between image data from different samples, cohorts or batches [25, 26].

Contouring
Tumors were segmented using a semi-automatic method via a gradient-based method (PET-Edge, MIM Software Inc), previously validated as a robust substitute of manual delineation for pancreatic tumors [19]. Ad hoc support of nuclear medicine physicians previously involved in radiomic investigations was available in case of difficult interpretation of the location of tumor on PET.

Extracting radiomic features

RF extraction was performed using the Spaarc Pipeline for Automated Analysis and Radiomics Computing (SPAARC), a software built on the MATLAB platform (MathWorks, Natick, MA, USA) [24,27], consistent with recently published IBSI guideline [25] for radiomic standardization; it utilises algorithms validated against all datasets from the IBSI international collaboration [25,26,28-30]. DICOM files were imported to MATLAB using the Computational Environment for Radiological Research (CERR) [31,32]. A discretization method at 64 fixed bin number was set in SPAARC, according to the suggestion by Tixier el al [33] and confirmed by ad-hoc tests [34]. One-hundred-ninety-eight RF of first and higher order were extracted, belonging to the families: Morphology, Statistical, Intensity Histogram, Grey Level Co-occurrence Matrix 3D_average (GLCM3D_avg), Grey Level Co-occurrence Matrix 3D_combined (GLCM3D_comb), Grey Level Run Length 3D_average (GLRL3D_avg), Grey Level Run Length 3D_combined (GLRL3D_comb), Grey Level Size Zone Matrix 3D, Neighbour Grey Tone Difference Matrix 3D (NGTDM3D), Grey Level Distance Zone Matrix 3D (GLDZM3D).

RF selection: robustness and cross-correlation

Two dedicated studies were previously carried out to assess the stability of RF against inter-observer variability during contouring and the impact of image acquisition/reconstruction parameters [19,34]. As our previous study on robustness with respect to acquisition/reconstruction parameters [34] was developed with a unique scanner, the Discovery-690, here the distributions of RF coming from images of patients scanned on the different scanners was investigated to assess any significant inter-scanner systematic variability. In total, 107, 39 and 33 patients were scanned on the Discovery-STE, the Discovery-ST and the Discovery-690 respectively. The impact of inter-scanner variability was assessed by the paired Mann-Whitney test to identify which RF extracted from images coming from different scanners belong to different distributions, and to exclude them in the RF selection process: a summary of these results is reported in the Supplementary material. In short, among 198 extracted, 91 RF were excluded as not robust with respect to inter-observer variability, based on intra-class correlation (ICC, with a 0.80 threshold) agreement tests of previously acquired multi-observer vs PET-edge contours [19]; among the remaining RF, 35/107 were excluded as likely suffering by acquisition/reconstruction parameters and inter-scanner variability.

The original population was randomly split in a training (n = 116) and a validation cohort (n = 60) according to the second level of the TRIPOD [35] guidelines for the validation of predictive models in oncology. In order to avoid redundancy, focusing on the training population, a correlation filter based on the Spearman correlation coefficient computation was applied on the remaining 72 RF, resulting in groups of highly cross-correlated RF (Spearman r >0.80). Then, among correlated RF, the ones with the minimum p-value were selected when testing their association with DRFS at univariate Cox regression.

At the end of this process, 7 RF (Morphology-COMshift; Statistical-percentile10 (P_{10}); GLSZM3D-small Zone Emphasis, small Distance Emphasis, grey Level Variance; NGTDM3D-complexity, high Dependence High Grey Level Emphasis) were then considered as the best candidates for multi-variate analysis. In Figure 1, a schematic plot of the RF selection procedure is shown.

In addition to RF, selected available clinical parameters were tested at multi-variate Cox regression: CA-19.9, sex, age, stage (IV vs III), prescribed radiotherapy dose, boost (yes/no) and tumor position (head vs others; body or body-tail vs others).
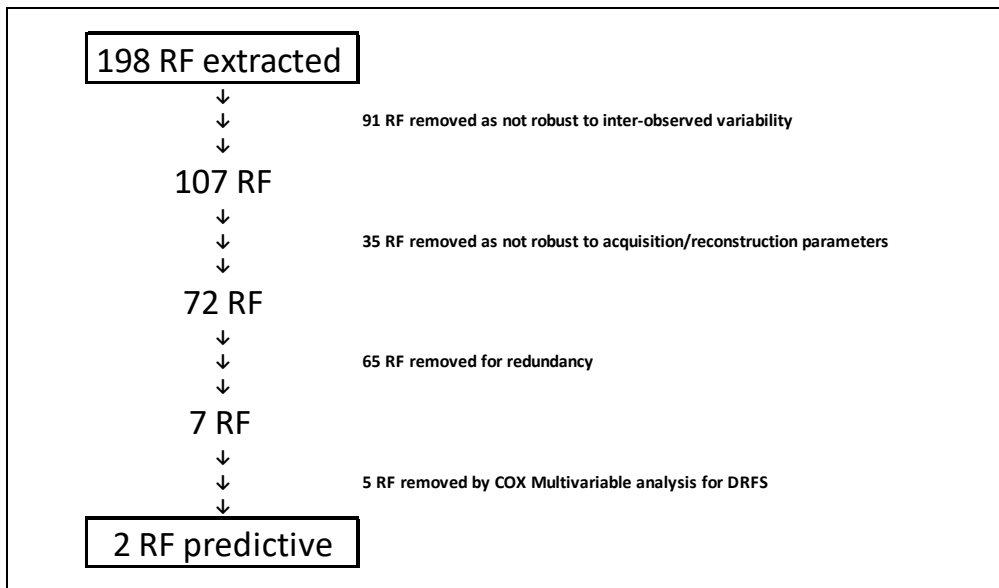
```
┌─────────────────────┐
│  198 RF extracted   │
└─────────────────────┘
         ↓
         ↓         91 RF removed as not robust to inter-observed variability
         ↓
      107 RF
         ↓
         ↓         35 RF removed as not robust to acquisition/reconstruction parameters
         ↓
       72 RF
         ↓
         ↓         65 RF removed for redundancy
         ↓
        7 RF
         ↓
         ↓         5 RF removed by COX Multivariable analysis for DRFS
         ↓
┌─────────────────────┐
│   2 RF predictive   │
└─────────────────────┘
```

Figure 1: Plot of the RF selection procedure

Statistics and modeling

Multi-variable analysis, considering the time from the beginning of RT and the endpoint occurrence, was performed on the training population combining the selected 7 RF in a Cox proportional hazard regression model with/without including clinical predictors for prediction of DRFS. A p-value < 0.05 and a backward selection was set to retain variables in the model. Prognostic indices (P_INDEX) were derived as the linear combination of the selected parameters weighted with their b value in the Cox model. The trained model for DRFS was then tested via Cox analysis on the validation cohort of 60 patients.

In order to represent the ability of the model in stratifying patients according to DRFS, DRFS curves of patients with the P_INDEX within the 1st quartile, higher than the 2nd quartile, and between the two quartiles were plotted and their separation was confirmed by the Kaplan-Meier test.

Sex (M/F)		85/91
Age (median; range) (y)		67 ; 44-86
Tumor volume (25th, 50th and 75th percentile) (cc)		7.35, 15.8, 32.68
Tumor site	body	25
	head	86
	tail	3
	head-body	9
	tail-body	20
	missing	33
Histology	Adenocarcinoma	154
	Cystoadenocarcinoma	10
	missing	12
Staging	III	158
	IV	17
PTV2 dose (Gy)		44.25 (30-58)
Patient with PTV2 boost		33
CA19.9 preRT (range; median)		1-6281; 112
Follow-up-median (range)		11 (1-37)

Table 1: Patients' characteristics

Results

DRFS Model –Training

The median follow-up from the start of CRT was 11 months with median time to distant relapse equal to 7 months (range: 1-33). The number of patients experiencing a distant relapse was 104. When combining the selected 7 RF in a Cox proportional hazard regression model for prediction of DRFS, only two parameters were selected backwardly: the Morphological COMshift and the Statistical P_{10} (p = 0.0005, HR = 2.72). Results are reported in Table 2. Values of DRFS P_INDEX ranged from -1.37 to 1.99, with a median of 0.49, and the 25th and 75th percentiles were equal to 0.28 and 0.76 respectively. To assess the ability of DRFS P_INDEX to properly represent different risk classes, patients were grouped in three subgroups: (1) patients with P_INDEX < 25th percentile, (2) > 75th percentile and (3) between 25th and 75th percentile. Kaplan Meier curves for DRFS, reported in Figure 2, showed an optimal ability to stratify risk classes (p = 0.0024). The P_INDEX was also applied on the entire population and the representation in risk classes was performed in the same way as in the training cohort, with Kaplan Meier curves, dividing patients in 3 groups based on quartiles. Cox results and Kaplan Meier curves are reported on Table S1 and Figure S2 of Supplementary Material.

DRFS Model –Validation

Cox analysis of the model including COMshift and P_{10} was repeated on the validation cohort. Results were reported in Table 2 and confirmed the ability of the model to predict DRFS (p = 0.03, HR = 2.53, 95%CL = 0.96-6.65). Very importantly, HR in the validation cohort was similar to the one of the training cohort. In the validation cohort the values of DRFS P_INDEX ranged from -2.23 to 1.37, with a median of 0.45, and the 25th and 75th percentiles equal to 0.26 and 0.75 respectively.

		TRAINING				VALIDATION			
	COX Covariate	b	HR	95%CL of HR	Overall fit Model p	COX Covariate	b	HR	95%CL of HR	Overall fit Model p
DRFS	Morphological COMshift	-0.3479	0.706	0.508 to 0.981	0.0022				
	Statistical Percentile10	0.0001	1.001	1.000 to 1.001					
	Prognostic index	1.0001	2.720	1.540 to 4.800	0.0005	Prognostic index	0.929	2.530	0.960 to 6.650	0.0296

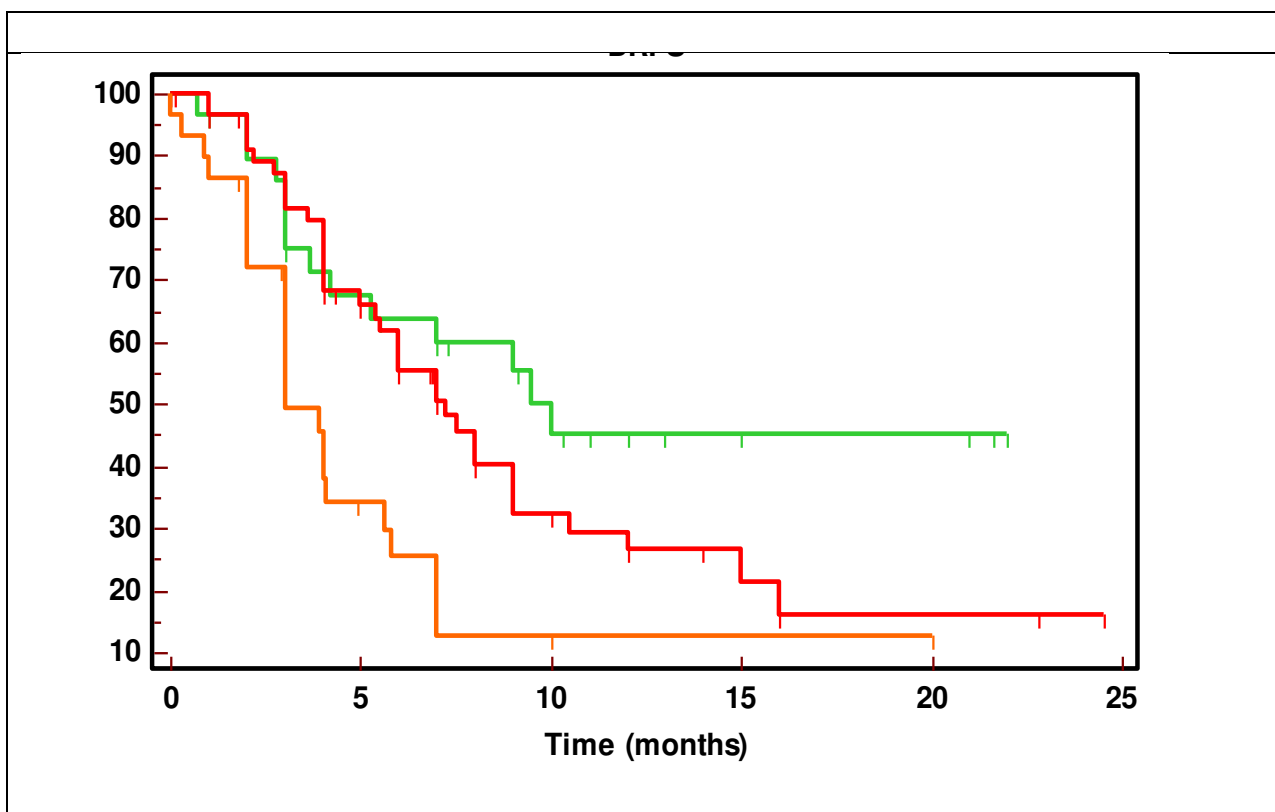*Table 2: Cox Analysis results for DRFS. Training and validation.*

Figure 2: Distant Relapse Free Survival probability curves according to P_INDEX values stratification for training cohort. The survival curves of patients with P_INDEX within the 1st quartile, higher than the 2nd quartile, and between the two quartiles are reported in green, orange and red respectively.

Impact of clinical variables

Multivariable Cox Analysis in the training cohort did not retain any independent impact of clinical variables, apart stage IV (*vs* III) was retained in the resulting model, with a relatively poor impact, as shown in the Supplementary material, mostly due to the quite small number of stage IV patients (n=17) compared to the total (n=176). Likely due to the same reason, the resulting model incorporating the two RF and "stage IV" was not fully confirmed in the validation cohort (p = 0.14).

Discussion

This study aimed to assess a robust radiomic signature able to predict DRFS based on data of 176 patients with LAPC treated with induction chemotherapy followed by chemo-radiotherapy, according to an institutional protocol, using moderate hypo-fractionation. The focus on distant relapse was justified by the fact that the development of distant metastatic disease represents the dominant pattern of tumor recurrence/progression for LAPC and is the most common cause of death. The possibility of selecting patients "a priori", on the basis of the predicted rapidity of the metastatic spread after CRT has the potential of dramatically improving the therapeutic approach.

In current study, PET images acquired before radiotherapy were used to extract the features which demonstrated the possibility to predict DRFS. We followed the second level approach of TRIPOD guideline, evaluating one random split of the cohort into training and validation sets. Our hypothesis was that a radiomics signature, if existing, should be found within the set of most robust and possibly interpretable features: a big effort was previously made to select RF robust with respect to items as the inter-observer variability during contouring and the impact of image acquisition/reconstruction parameters. After the redundancy analysis, a robust predictive model based on two independently predictive RF was developed. To our knowledge this is the first case of study on LAPC disease extended on a large population, using a validated semi-automatic segmentation method and an IBSI consistent software for radiomic features calculation and extraction, suggesting a robust, internally validated, model with few variables. Other similar studies showed possible lacks in robustness [18] and risks of overfitting due to the limited number of patients and the relative high number of variables included [16,17,36].

It is well known that the large majority of LAPC patients develops distant relapse, and the early identification of the group with "low" probability of early metastases via radiomic signature could permit: (1) to pave the way to new trials, intensifying local and/or systemic therapies for patients with better prognosis, with a potentially relevant improvement of the therapeutic outcome and (2) to avoid over-treatment of patients with expected poor outcome.

Intriguingly, few considerations could be attempted regarding the possible "biological" meaning of the two predictive RF: P_{10} is an intensity-based statistical feature which represents the 10th percentile of the set of grey levels of the voxels included in the ROI segmented as PET positive lesion. Grey levels in PET represent the uptake intensity, and in our population high values of P_{10}, increasing the P_INDEX, were found to be associated to patients more prone to early develop distant metastasis. If P_{10} is large among the grey level set of the ROI, the 90% of intensities in the ROI are even larger, that is compliant for lesions with a deep uptake, showing a "compact" aspect, with little blurring at the edges. Also, the corresponding intensity histograms are characterized by a shape more shifted versus higher intensities.

Regarding the COMshift belonging to the Morphological family of RF, it is the distance between the ROI centroid and the intensity-weighted ROI centroid. The larger is the shift, the larger are the heterogeneity within the ROI segmented. On the contrary small shifts represent lesions with a more homogeneous distribution of uptake, without discontinuities and internal regions of low uptake. In that way, small values of COMshift contribute to preserve a high P_INDEX value, namely a higher risk of early metastasis. Two examples are shown in Figure 3: in the upper part, the image of a patient without distant relapse at the time of death is reported. In the coronal and sagittal images, the lesion uptake is not distributed uniformly in the ROI, showing sub-regions of low uptake. In this case the COMshift is high and so rightly associated by the model with low risk of early metastasis: for this case, the low value of P_{10} has less impact on the prognostic index value. The lower part of Figure 3 shows a patient with early distant relapse. In the coronal and sagittal images, the lesion uptake is highly concentrated and quite uniformly distributed in the ROI. In this case the COMshift

is close to zero. The P_INDEX value is determined by both the high value of P_{10} and the small COMshift, classifying this case in the high-risk class.

Only the paper by Cui et al. [18], may reasonably be compared against our study as they followed in part the same second-level TRIPOD approach, rigorous methods to avoid overfit during variable selection and Cox regression in a sufficiently large population of LAPC patients. Their study was however focused on overall survival and patients were treated with SBRT: especially the different end-point makes impossible any strict comparison. Of note, the authors found a 7-variable radiomic index to predict OS, mostly related to intra-tumor heterogeneity, in part consistent with the finding regarding COMshift. Differently from our study, in Cui et al. GTV was delineated on CT and overlayed on PET without any PET segmentation; moreover, the impact of inter-observer variability was simply assessed by rigidly modifying contours and looking to the changes induced in the features, which is a rough approximation of the true inter-observer variability reported for TC [37]. In our opinion, this method, together with the no-IBSI compliance of the used software at that time, makes their findings poorly generalizable. Instead, our approach to apply a validated semi-automatic segmentation method using a commercial software and an IBSI-consistent software for RF extraction, as well as a correct evaluation of the impact of the (large) inter-observer variability should make our results more usable by the community.

More in general, the generalizability of the proposed index may be expected to be high also due to the very small number of robust features (only two) and to their characteristics: the absence in the index of high-order features should be considered positively from the point of view of inter-scanner/inter-center variability of the suggested index.

A potential pitfall regards the retrospective nature of the study and the heterogeneity of patient's characteristics: in particular, 17/176 patients were already metastatic at the disease onset: we decided to retain them in this cohort study as they received the same treatment of stage III patients, (being under control after induction chemotherapy), considering as endpoint the occurrence of distant relapse in the first new metastatic site. Of note, if repeating the analyses on the whole group with/without including them, the two selected RF remained independently predictive.

Regarding the impact of clinical variables, only stage IV resulted to be associated to early distant relapse in the training cohort, not fully confirmed in the validation one. Surely it is a matter of small numbers, since that when considering all the entire population, the P_INDEX generated by COMshift, P_{10} and stage resulted to be highly predictive, confirming that a larger validation cohort would be needed to corroborate the inclusion of the stage in the predictive model.

Within the limitations of the study, we are confident that this model could work also in other LAPC populations, therefore is intended to be externally validated: a multi-centric study with this aim, involving three new Institutes, is under activation.

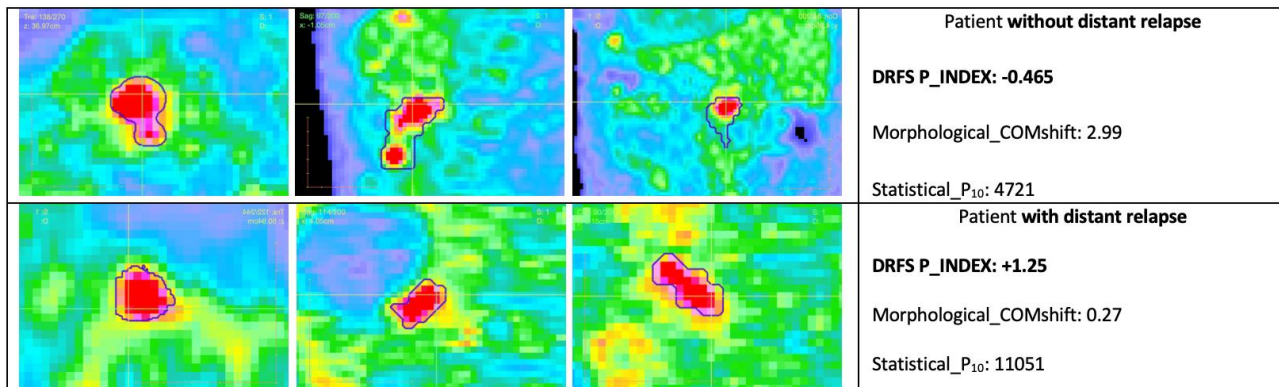|	Patient **without distant relapse**
	DRFS P_INDEX: -0.465
	Morphological_COMshift: 2.99
	Statistical_P_{10}: 4721
	Patient **with distant relapse**
	DRFS P_INDEX: +1.25
	Morphological_COMshift: 0.27
	Statistical_P_{10}: 11051

Figure 3: In the upper part a patient without distant relapse at the time of death is reported. The lesion uptake is not distributed uniformly in the ROI, showing sub-regions of low uptake. In this case the COMshift is high and so rightly associated by the model with low risk of early metastasis: for this case, the low value of P_{10} has less impact on the prognostic index value. The lower part a patient with very early distant relapse is reported. The lesion uptake is highly concentrated and uniformly distributed in the ROI. In this case the COMshift is close to zero. The P_INDEX value is determined by both the very high value of P_{10} and the small COMshift, classifying this case to the high-risk class.

Conclusions

The results referred to a large cohort of inoperable LAPC patients showed that robust PET RF may predict DRFS after CRT with a more than promising evidence of being able to identify risk classes in a substantially improved manner compared to available biomarkers/clinical features. A fair discrimination power was found applying the model to training and validation samples. Further validation studies would be advisable to confirm these findings.

References:

1. Siegel R, Naishadham D, Jemal A. Cancer statistics, CA Cancer J Clin. 2013;63:11-30.

2. WHO report on cancer, 2020, Geneve.

3. Vincent A, Herman J, Schulick R. Pancreatic cancer. Lancet. 2011;378:607-20.

4. American Cancer Society. Cancer Facts & Figures 2020. Atlanta: American Cancer Society; 2020. https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2020/cancer-facts-and-figures-2020.pdf

5. Ben-Josef E, Schipper M, Francis I et al. A phase I/II trial of intensity modulated radiation (IMRT) dose escalation with concurrent fixed-dose rate gemcitabine (FDR-G) in patients with unresectable pancreatic cancer. Int J Radiat Oncol Biol Phys. 2012;84:1166-1171.

6. Berger AC, Garcia M Jr, Hoffman JP et al. Postresection CA 19-9 predicts overall survival in patients with pancreatic cancer treated with adjuvant chemoradiation: a prospective validation by RTOG 9704. J Clin Oncol 2008;26:5918-5922.

7. Fong ZV, Winter JM. Biomarkers in pancreatic cancer: diagnostic, prognostic, and predictive. Cancer J. 2012;18:530-538.

8. Harsha HC, Kandasamy K, Ranganathan P et al. A compendium of potential biomarkers of pancreatic cancer. PLoS Med 2009;6:e1000046-e1000046.

9. Hidalgo M. Pancreatic cancer. N Engl J Med. 2010;362:1605-1617.

10. Incerti E, Vanoli EG, Broggi S et al. Early variation of 18-fluorine-labelled fluorodeoxyglucose PET-derived parameters after chemoradiotherapy as predictors of survival in locally advanced pancreatic carcinoma patients. Nucl Med Commun. 2019 Oct;40(10):1072-1080.

11. Wilson JM, Mukherjee S, Brunner TB et al. Correlation of 18F-Fluorodeoxyglucose Positron Emission Tomography Parameters with Patterns of Disease Progression in Locally Advanced Pancreatic Cancer after Definitive Chemoradiotherapy Clinical Oncology 2017;29(6):370-377.

12. Mellon EA, Jin WH, Frakes JM et al. Predictors and survival for pathologic tumor response grade in borderline resectable and locally advanced pancreatic cancer treated with induction chemotherapy and neoadjuvant stereotactic body radiotherapy. Acta Oncologica 2017;56(3):391-397.

13. Lambin, P. et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur. J. Cancer 2012; 48:441–446.

14. Konert T, Everitt S, La Fontaine MD, et al. Robust, independent and relevant prognostic 18F-fluorodeoxyglucose positron emission tomography radiomics features in non-small cell lung cancer: Are there any?. *PLoS One*. 2020;15(2).

15. Zhu A, Lee D , Shim H. Metabolic positron emission tomography imaging in cancer detection and therapy response. Semin Oncol 2011;38:55-69.

16. Yue Y, Osipov A, Fraass B et al. Identifying prognostic intratumor heterogeneity using pre- and post-radiotherapy 18F-FDG PET images for pancreatic cancer patients. Journal of Gastrointestinal Oncology 2017;8(1):127-138.

17. Yue Y, Yang W, Fraass B et al. Prognostic Modeling of Locally Advanced Pancreatic Cancer Treated With Radiation Therapy Using [18F] FDG-PET Features and CA-19-9. Int J Radiat Oncol Biol Phys 2014;90:S361.

18. Cui Y, Song J, Pollom E et al.Quantitative Analysis of 18F-Fluorodeoxyglucose Positron Emission Tomography Identifies Novel Prognostic Imaging Biomarkers in Locally Advanced Pancreatic Cancer Patients Treated With Stereotactic Body Radiation Therapy. International Journal of Radiation Oncology Biology Physics 2016;96(1):102-109.

19. Belli ML, Mori M, Broggi S et al. Quantifying the robustness of [^{18}F]FDG-PET/CT radiomic features with respect to tumor delineation in head and neck and pancreatic cancer patients. Phys Med 2018; 49:105-111.

20. Shah MH, Goldner WS, Halfdanarson TR *et al*. NCCN guidelines insights: neuroendocrine and adrenal tumors, version 2. *J Natl Compr Canc Netw* 2018; 16:693–702.

21. Cattaneo MG, Passoni P, Longobardi B et al. Dosimetric and clinical predictors of toxicity following combined chemotherapy and moderately hypofractionated rotational radiotherapy of locally advanced pancreatic adenocarcinoma. Radiother. Oncol 2013;108(1):66-71.

22. Passoni P, Reni M, Cattaneo GM et al. Hypofractionated image-guided IMRT in advanced pancreatic cancer with simultaneous integrated boost to infiltrated vessels concomitant with capecitabine: a phase I study. Int J Radiat Oncol Biol Phys 2013;87:1000-1006.

23. Reni M, Cereda S, Rognone A et al. A randomized phase II trial of two different 4-drug combinations in advanced pancreatic adenocarcinoma: cisplatin, capecitabine, gemcitabine plus either epirubicin or docetaxel (PEXG or PDXG regimen). Cancer Chemother Pharmacol 2012;69:115-23.

24. Whybra P, Parkinson C, Foley K et al. Assessing radiomic feature robustness to interpolation in (18) F-FDG PET imaging. Sci Rep. 2019;9(1):9649.

25. Zwanenburg A, Vallières M, Abdalah MA, et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. Radiology. 2020;00:1-11

26. Zwanenburg A, Leger S, Vallières M et al. Initiative, for the I. B. S. Image biomarker standardisation initiative. *arXiv:1612.07003* (2016).

27. Gwynne S, Spezi E, Wills L et al. Toward semi-automated assessment of target volume delineation in radiotherapy trials: the SCOPE 1 pretrial test case. Int J Radiat Oncol Biol Phys. 2012;84(4):1037-42.

28. Zwanenburg A et al. PO-0981: Results from the Image Biomarker Standardisation Initiative. Radiother. Oncol. 2018;127:S543–S544.

29. Parkinson C et al. Evaluation of prognostic models developed using standardised image features from different PET automated segmentation methods. EJNMMI Res. 2018;8:29.

30. Whybra P, Foley K, Parkinson C et al. Effect of Interpolation on 3D Texture Analysis of PET Imaging in Oesophageal Cancer. Radiother Oncol. 2018;127(Supp 1):S1167-1168.

31. Deasy JO, Blanco AI, Clark VH. CERR: A computational environment for radiotherapy research. Med. Phys. 2003;30:979-985.

32. Apte AP et al. Technical Note: Extension of CERR for computational radiomics: A comprehensive MATLAB platform for reproducible radiomics research. Med. Phys. 2018;45:3713-3720.

33. Tixier F, Hatt M, Le Rest CC et al. Reproducibility of tumor uptake heterogeneity characterization through textural feature analysis in 18F-FDG PET. *J Nucl Med*. 2012;53(5):693-700.

34. Presotto L, Bettinardi V, De Bernardi E et al. PET textural features stability and pattern discrimination power for radiomics analysis: An "ad-hoc" phantoms study. Phys Med 2018;50:66-74.

35. Collins GS, Reitsma JB, Altman DG et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. Br J Cancer 2015;112(2):251-259.

36. Tuli R, Fraass B, Yang W, et al. Pretreatment 18 F-FDG-PET texture analysis to predict local response of pancreatic cancer to radiotherapy. J Clin Oncol 2014;32:375.

37. Caravatta L, Cellini F, Simoni N, et al. Magnetic resonance imaging (MRI) compared with computed tomography (CT) for interobserver agreement of gross tumor volume delineation in pancreatic cancer: a multi-institutional contouring study on behalf of the AIRO group for gastrointestinal cancers. Acta Oncologica 2019;58(4):439rt-447.