Contents lists available at ScienceDirect

Physics and Imaging in Radiation Oncology

journal homepage: www.elsevier.com/locate/phro

phiRO
Physics and Imaging in Radiation Oncology

Original Research Article

Combined morphologic and metabolic pipeline for Positron emission tomography/computed tomography based radiotherapy response evaluation in locally advanced pancreatic adenocarcinoma

Yi Lao[a], John David[a], Arman Torosian[a], Veronica Placencio[a,b], Yalin Wang[b], Andrew Hendifar[c], Wensha Yang[a], Richard Tuli[a,*]

[a] Department of Radiation Oncology, Cedars-Sinai Medical Center, Los Angeles, USA
[b] School of Computing, Informatics, Decision Systems and Engineering, Arizona State University, Tempe, AZ, USA
[c] Department of Surgery, Cedars-Sinai Medical Center, Los Angeles, USA

ABSTRACT

Background and purpose: Adaptive radiation planning for pancreatic adenocarcinoma (PA) relies on accurate treatment response assessment, while traditional response evaluation criteria inefficiently characterize tumors with complex morphological features or intrinsically low metabolism. To better assess treatment response of PA, we quantify and compare regional morphological and metabolic features of the 3D pre- and post-radiation therapy (RT) tumor models.
Materials and methods: Thirty-one PA patients with pre and post-RT Positron emission tomography/computed tomography (PET/CT) scans were evaluated. 3D meshes of pre- and post-RT tumors were generated and registered to establish vertex-wise correspondence. To assess tumor response, Mahalanobis distances ($M_{dist}|_{Fusion}$) between pre- and post-RT tumor surfaces with anatomic and metabolic fused vectors were calculated for each patient. $M_{dist}|_{Fusion}$ was evaluated by overall survival (OS) prediction and survival risk classification. As a comparison, the same analyses were conducted on traditional imaging/physiological predictors, and distances measurements based on metabolic and morphological features only.
Results: Among all the imaging/physiological parameters, $M_{dist}|_{Fusion}$ was shown to be the best predictor of OS (HR = 0.52, p = 0.008), while other parameters failed to reach significance. Moreover, $M_{dist}|_{Fusion}$ outperformed traditional morphologic and metabolic measurements in patient risk stratification, either alone (HR = 11.51, p < 0.001) or combined with age (HR = 9.04, p < 0.001).
Conclusions: We introduced a PET/CT-based novel morphologic and metabolic pipeline for response evaluation in locally advanced PA. The fused $M_{dist}|_{Fusion}$ outperformed traditional morphologic, metabolic, and physiological measurements in OS prediction and risk stratification. The novel fusion model may serve as a new imaging-marker to more accurately characterize the heterogeneous tumor RT response.

1. Introduction

Pancreatic adenocarcinoma (PA) has one of the worst outcomes of all solid tumors with a 5-year survival rate of 7% [1,2]. At the time of diagnosis, only 10% − 15% of patients are eligible for resection, which is the only therapy offering potentially prolonged survival [1]. A major contraindication for surgery is vascular involvement by primary tumor, which precludes a margin-negative resection. As a result, approximately 40% are considered borderline resectable or locally advanced (BRPA, LAPA). Following aggressive multiagent chemotherapy and radiotherapy, only < 30% of patients are downstaged [3]. However, an

accurate post-radiotherapy assessment with computed tomography (CT) remains challenging and may not accurately reflect tumor response [4].

To assess tumor treatment response, size-based criteria, such as Response evaluation criteria in solid tumors (RECIST) and World Health Organization (WHO), have been widely adopted in current clinical practice [5,6]. While easy to implement and generalize, these uni-dimensional measurements inaccurately characterize tumors with complex geometry (Fig. 1 (A)) and are insensitive to tumor heterogeneity, which is particularly important for dense stromal tumors such as PA [7]. For example, in a study investigating the association between
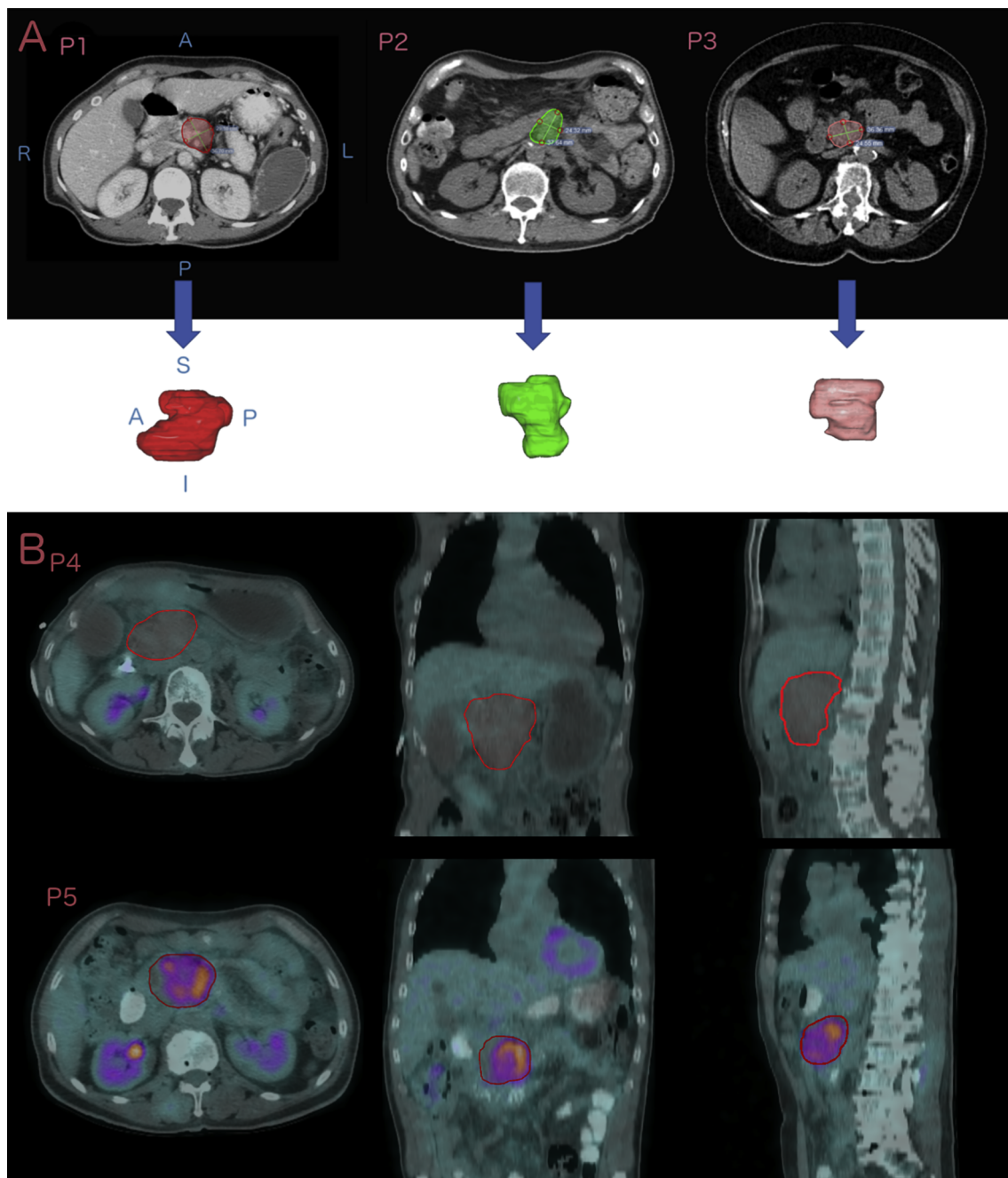

Fig. 1. A: Representative tumor short and long axes on axial CT images and the corresponding 3D constructions for a 64-year-old female (P1), a 76-year-old male (P2) and a 77-year-old female pancreatic adenocarcinoma (PA) patient (P3). All 3 patients show similar values of tumor short and long axes [5,6], while their 3D tumor constructions revealed varied shape features. B: Representative PET/CT fusion images of PA patients with varied SUV distributions within tumors. Boundaries of tumors are circled in red. P4, Axial (left column), coronal (middle column), and sagittal (right column) views of a 77-year-old female patient with low SUV PA tumor. P5, a 61 years old male patient with heterogeneously distributed SUV within PA tumor. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

RECIST response and overall survival (OS), similar median OS durations were observed in patients who underwent resection and had no radiographic evidence of response to neoadjuvant therapy and patients who did respond [4]. In addition to size-based criteria, F Fluorodeoxyglucose (FDG) positron emission tomography (PET) has been commonly used to quantify tumor metabolic activities [8,9,10]. However, basic PET values, such as the maximum or mean FDG uptake, is inadequate to describe the heterogeneous tumor metabolism (Fig. 1 (B), P5).

Moreover, not all tumors present with increased FDG uptake compared to surrounding normal parenchyma [8] and comorbidities such

as hyperglycemia and diabetes, which are common in PA, can disrupt tissue FDG uptake leading to false-negative results [11]. Therefore, traditional assessments based on current standard of care functional (SUVmax, SUVmean, etc.) and structural imaging (RECIST, WHO criteria) are inadequate to evaluate treatment response in PA. Integrated anatomic and metabolic assessment may provide a more comprehensive and accurate indication of tumor response [10]. In this study, we integrate morphologic and metabolic metrics of LAPA and BRPA using CT/PET to assess treatment response and clinical outcomes using multivariate distances between the resulted 3D pre- and post- RT tumor surfaces. Our aims are: first, to develop a combined 3D anatomic and

metabolic pipeline; second, to test the efficacy of this pipeline as a quantitative survival predictor for patients with LAPA or BDPA.

2. Materials and methods

2.1. Subjects and data

With Institutional Review Board approval, 31 patients with biopsy-confirmed PA were retrospectively selected from our institutional database between 2011 and 2017. All patients had LAPA/BRPA treated with radiotherapy and received both pre and post-RT PET with diagnostic CT scans. The post-RT PET/CT was taken 6–8 weeks after radiotherapy.

Patients were instructed to fast for a minimum of 6 h and then undergone skull to thigh imaging with a PET/CT scanner. Three scanners were included in this study: one Discovery 710 (64-slice PET/CT, General Electric Healthcare, Boston, Massachusetts, U.S), one Biograph (64-slice TruePoint PET/CT, Siemens Healthcare, Erlangen, Germany) and one Gemini (64-slice PET/CT, Philips Healthcare, Amsterdam, Netherlands). All PET/CT scans were obtained approximately 50 mins after injection of F-FDG ($5\,MBq/kg$). All the acquired PET images were constructed using CT-based attenuation correction. Quality assurance was routinely performed to maintain consistent calibration on all PET/CT. In addition, all patients had carbohydrate antigen 19-9 (CA 19-9) values – the most extensively studied and validated physiological biomarker in PA prognosis [12] - measured pre-RT, and twenty-nine of them had CA 19-9 collected post-RT.

2.2. Data preprocessing

On attenuation corrected PET scans, standardized uptake values (SUV) were calculated and normalized to the patient's body weight as:

$$SUV_{BW} = \frac{C_{img}(\mathrm{T})}{D_A/BW} \tag{1}$$

where $C_{img}(T)$ is the radioactivity concentration measured at time T, D_A denotes administrated dose [MBq] at the time of injection, and BW represents patient's body weight [kg] on the day of imaging. For simplicity, SUV_{BW} will be shorted as SUV in the remaining text. To transform the PET indices into the CT space, SUV maps were up-sampled to the resolutions of respective CT images and used as the main source of metabolic inputs in later steps.

For each patient, boundaries of baseline and residual primary pancreatic gross tumor volumes (GTV) were manually defined by a single experienced gastrointestinal radiation oncologist on pre- and post-RT diagnostic CT images. To assist tumor definition, the contrast CT scans were registered to the corresponding PET/CT using a commercial B-spline based deformable image registration software (Velocity AI™ Varian, Palo Alto, CA). The intra-observer reproducibility was evaluated with the same radiation oncologist contouring the tumor twice for four patients at two different time points eight weeks apart. The mean Jaccard index (intersection over union) was used to compare intra-observer variability. Pre- and post-RT CT images were rigidly registered to align the tumor orientation. The resultant transformation matrices were propagated to the corresponding PET images and tumor segmentation. This step enables following regional tumor surface matching. The propagated pre- and post-RT tumor segmentations (GTV$_{pre}$ and GTV$_{post}$) were binarized and assessed using the pipeline illustrated in Fig. 2.

2.3. Surface measurement generation

As shown in Fig. 2, a 3D surface mesh was generated from GTV segmentation using the matching cube algorithm [13]. To obtain a regional correlation of pre- and post-RT tumor surfaces, GTV$_{pre}$ and

GTV$_{post}$ were matched using surface-based constrained harmonic registration [14,15]. As such, shape changes were tracked within the structure-of-interest, excluding the influence of surrounding parenchyma [16]. For each surface vertex, a multivariate vector was formed using both the anatomic and metabolic components.

The anatomic component consisted of two geometrical features based on the same surface model: the normal direction, represented by the midline to surface radius ρ, and the surface area, represented by the Jacobian of the transformation (J) from surface registration [16,17]. Two Jacobian derived metrics were used: the determinant of Jacobian (det$J$) and the log-Euclidean projected deformation tensor ($\log\sqrt{JJ^T}$). det$J$ indicates the extent of warping needed to deform a surface to match its target, and $\log\sqrt{JJ^T}$ further encapsulates directional information in the deformation [15,16].

To characterize the surface vertices using tumor metabolism, the midlines of all the 3D tumor volumes were first determined. In CT-aligned SUV maps, SUV measurements within the 3D tumor mesh were then sampled as surface vertices along the corresponding radius to midline, as illustrated in [19]. Specifically, using [18,19]:

$$\left\| \frac{(\overrightarrow{X-M}) \times (\overrightarrow{P-M})}{\|\overrightarrow{X-M}\|} \right\| \ll D \tag{2}$$

and

$$(\overrightarrow{X-P}) \bullet (\overrightarrow{P-M}) \geq 0 \tag{3}$$

where X, M, P denote the (x, y, z) coordinates of a vertex in the surface, the corresponding point of the vertex in the midline, and a voxel within the 3D representations, respectively. A pre-defined distance ($D = 1.2$) were chosen to ensure necessary voxel assignment while minimizing the overlap with neighboring vertices.

2.4. Statistical analysis

Paired Mahalanobis distance (M_{dist}), a dissimilarity measurement of observations in multivariate space, was used to assess tumor response. Specifically, for each tumor, given p-dimensional pre-RT surface observation $X_{prei} = [X_{pre1}, X_{pre2}, \cdots X_{prep}]^T$, and the corresponding post-RT observation $X_{posti} = [X_{post1}, X_{post2}, \cdots X_{postp}]^T$, the sample covariance matrix was calculated as:

$$\bar{S} = \frac{1}{n-1} \sum_{i=1}^{n} (X_{di} - \bar{X}_d)(X_{di} - \bar{X}_d)' \tag{4}$$

where n is the number of surface vertices, X_d denotes the difference between X_{pre} and X_{post}, and $\bar{X}_d$ represents the arithmetic mean vector of X_d over the surface vertices.

For each patient, M_{dist} of pre- and post- RT surfaces was obtained using:

$$M_{dist} = \bar{X}_d \bar{S}^{-1} \bar{X}_d' \tag{5}$$

Surface variables were analyzed accordingly: 1) metabolic features only - the surface projected SUV_{max} and SUV_{mean}; 2) morphological features only – radius ρ and $\log\sqrt{JJ^T}$; 3) combined metabolic and morphological features – ρ, det$J$, and SUV_{max}. While $\log\sqrt{JJ^T}$ adds directional information in surface warping, it also increases the noise when combined with other metabolic and morphological features [16]. As a result, we chose det$J$ over $\log\sqrt{JJ^T}$ in the combined metrics. M_{dist}s were calculated based on these three combinations and denoted as $M_{dist}|_{SUV}$, $M_{dist}|_{Shape}$, and $M_{dist}|_{Fusion}$, respectively.

The prognostic values of the three M_{dist} metrics were evaluated using Cox proportional hazards regression with OS. OS was defined as the time span from the time of diagnosis to the time of patient death or last encounter. Correlation analyses were performed between OS and baseline (B) measurements of serum CA19-9, GTV, global SUV_{max}, global SUV_{mean}, as well as post-RT versus pre-RT ratios (R) of the above
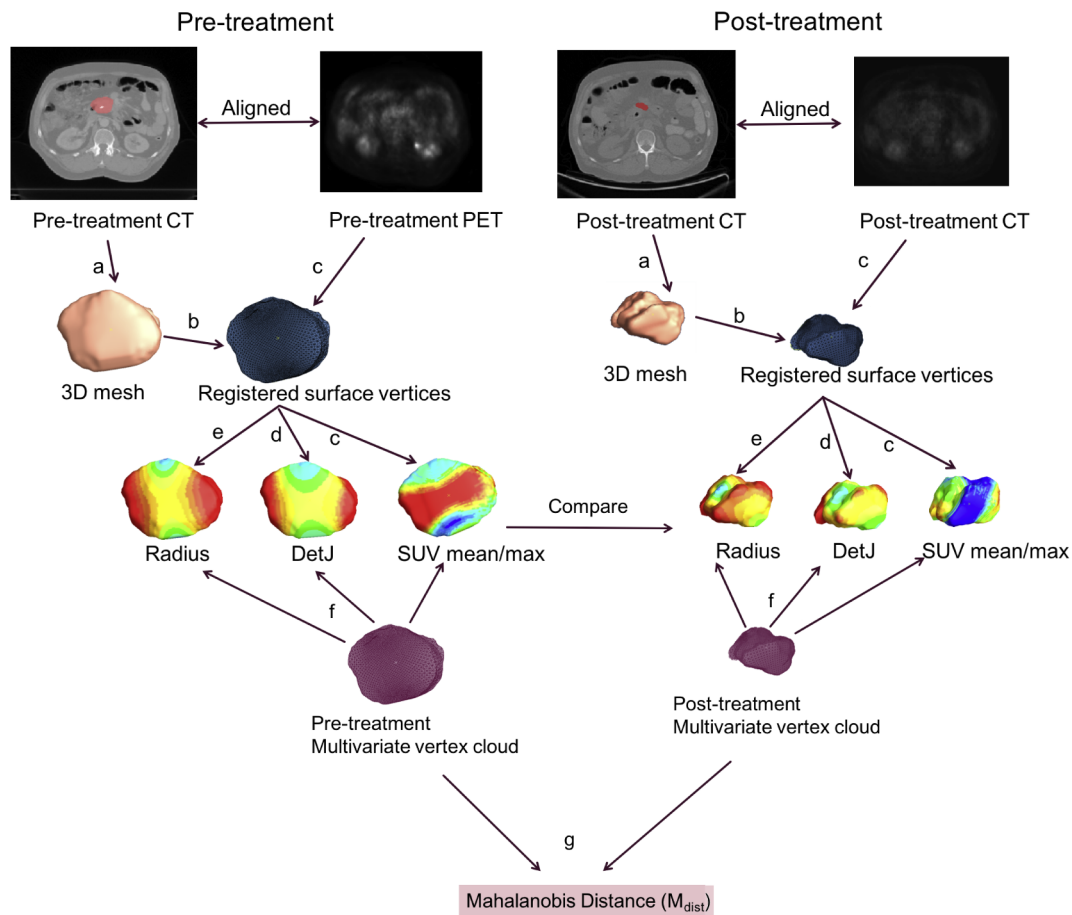
Fig. 2. Illustration of our pipeline. Processing steps represented in numbers are: a) 3D construction; b) Surface-based registration; c) PET indices sampling; d) Anatomic parameter (det*J*) calculation; e) Anatomic parameter (Radius) calculation; f) Multivariate surface vertex cloud assembling; g) Mahalanobis distance calculation on multivariate surface vertex ouds.

four measurements. The enrolled subjects in our study were gender-matched (52% male and 48% female) but had a wide range of age. As a result, we also investigated the prognostic value of age as an OS predictor.

To evaluate the classification capacity of each risk predictor, all patients were stratified into two risk groups based on the three M_{dist} metrics. The low versus high-risk populations were determined according to 2 methods, M1 and M2. M1 first sorted each of the above metrics and then grouped the lower half subjects as a low-risk group (LG) and the upper half subjects as high-risk group (HG). M2 used k-means clustering algorithm to automatically separate LG and HG based on multivariate data, consisting of the tested metric and patient's age at diagnosis. The inclusion of age in the analysis was again due to the wide range of age presented in this study and the significant association between age and OS, as revealed by our data. Log-rank tests were performed between LG and HG stratified patients using both M1 and M2, respectively. As a comparison, log-rank statistics of M_{dist} derived metrics were also compared with widely used imaging metrics (post- to pre-RT ratios of global SUV_{max}, global SUV_{mean}, and GTV). Additionally, because the enrolled subjects were heterogeneous in terms of gender, resectability (BRPA versus LAPA), RT methods (SBRT versus CRT), and post-treatment resection (unresected versus resected), survival differences between subgroups were also examined by log-rank tests.

3. Results

Patient demographics and disease characteristics are listed in Table 1. The age of the enrolled subjects ranged from 45 to 84 years

with a mean of 69.2 ± 9.5 (standard deviation) years. The mean Jaccard index of intra-observer GTV contour variability was 0.89. Results of univariate Cox regression analyses of all predictors are shown in Table 2. Among all the imaging parameters, $M_{dist}|_{Fusion}$ - the extent of tumor response considering both the morphologic and metabolic image information, was the best predictor of OS ($p = 0.008$) with a hazard ratio (HR) of 0.52 (95% CI, 0.19–0.78). Post-RT versus pre-RT ratio of SUV_{max} was associated with OS with borderline significance (HR = 3.89, 95% CI, 0.94–16.02, $p = 0.060$). All other imaging metrics (baseline or ratios of global SUV_{mean}, GTV, as well as M_{dist} metrics that based on morphologic or metabolic information alone) were not associated with OS. Additionally, patient's age at diagnosis was significantly associated with OS, (HR = 1.14, 95% CI, 1.04–1.25, $p = 0.004$), while traditional physiological biomarker CA19-9 did not present significant association with OS in our small cohort of patients.

The capability of each measurement to classify patients is more evidently reflected in risk stratification analysis. Patients presented no survival differences between genders (female vs. male, $p > 0.1$), resectability (BRPA vs. LAPA, $p > 0.1$), treatment methods (SBRT vs. CRT, $p > 0.1$), and post-treatment surgery (unresected vs. resected, $p > 0.1$). Kaplan-Meier plots of risk groups stratified using M1 and M2 are shown in Fig. 3, respectively. Results of log-rank tests comparing each pair of the survival curves can be found in Table 3. Evidently, from Fig. 3, the survival risk of patients was best stratified by $M_{dist}|_{Fusion}$, either alone or combined with patients' age. Using sorted levels of each parameter (M1), $M_{dist}|_{Fusion}$ successfully classified an HG group with a mean of 0.57, and LG with a mean of 4.65, the former has a hazard ratio almost twelve times higher than the latter ($p < 0.001$). $M_{dist}|_{SUV}$, the

Table 1
Patient characteristics.

Characteristics	Patients (%) (N = 31)
Gender	
Male	16 (52%)
Female	15 (48%)
Age at Baseline Scan	
Median	69
Mean	69.2
Std	9.5
Range	45–84
Clinical Tumor Stage	
Stage I	1 (3%)
Stage II	1 (3%)
Stage III	7 (23%)
Stage IV	21 (68%)
Unknown	1 (3%)
NCCN[1] Resectability	
Locally Advanced	23 (74%)
Borderline Resectable	8 (26%)
Primary Tumor Location	
Uncinate, head, and neck	23 (74%)
Body and tail	4 (13%)
Combined	4 (13%)
Scan and Treatment Time Interval	
Pretreatment (w)	4.54 (3.76)
Posttreatment (w)	7.14 (4.28)
RT Type[2]	
SBRT	6 (19%)
CRT	25 (81%)
Resected[3]	
Yes	11 (35%)
No	20 (65%)

[1] National Comprehensive Cancer network.
[2] RT types include stereotactic body radiation therapy (SBRT) and chemoradiotherapy (CRT).
[3] This refers to the resection after RT.

Table 2
Results of univariate Cox proportional hazards regression between OS and 12 predictors. Four traditional prognostic measurements (serum CA19-9, global SUV_{max}, global SUV_{mean}, and GTV) are categorized as baseline measurement (B), and post-RT versus pre-RT ratio (R). Surface measurements based on metabolic features only, anatomic features only and combined anatomic and metabolic features are indicated as $M_{dist}|_{SUV}$, $M_{dist}|_{Shape}$, $M_{dist}|_{Fusion}$, respectively. Statistical significance are marked as * for *p*-values < 0.05, and ** for *p*-values < 0.01.

	Mean	Std	HR/exp(β) (95% CI)	*p* values
Age	68.20	9.52	1.14 (1.04–1.25)	0.004**
CA19-9 (B) (k)	2.62	11.8	0.99 (0.99–1.00)	> 0.1
CA19-9 (R) (n = 29)	0.68	0.66	1.63 (0.70–3.77)	> 0.1
SUV_{max} (B)	6.00	2.37	0.90 (0.72–1.13)	> 0.1
SUV_{max} (R)	0.81	0.39	3.89 (0.94–16.02)	0.060
SUV_{mean} (B)	2.41	0.63	0.81 (0.32–2.04)	> 0.1
SUV_{mean} (R)	0.91	0.31	1.75 (0.45–6.75)	> 0.1
GTV (B) (*cc*)	51.71	35.60	1.00 (1.00–1.00)	> 0.1
GTV (R)	0.90	0.38	3.01 (0.57–15.71)	> 0.1
$M_{dist}|_{SUV}$	3.25	4.87	0.39 (0.24–1.13)	> 0.1
$M_{dist}|_{Shape}$	0.70	0.77	1.22 (0.69–2.14)	> 0.1
$M_{dist}|_{Fusion}$	2.68	4.62	0.52 (0.19–0.78)	0.008**

surface based distance measurement based on metabolic features only, also classified two groups of patients with significantly different risks ($p = 0.021$), while all other parameters failed in classifying survival risks using M1. Using unsupervised clustering (M2), a combination of $M_{dist}|_{Fusion}$, and age significantly partitioned two risk groups ($p < 0.001$), with the HG having a nine-fold hazard risk than the LG. M_{dist} calculated on metabolic features alone ($M_{dist}|_{SUV}$), and post- to pre-RT ratio of SUV_{max} also significantly classify HG and LG, with respective *p* values equal to 0.015, and 0.008. Other combinations were unable to stratify the patient cohort.

4. Discussion

Accurate response assessment during and after radiotherapy for LAPA and BRPA is critical to clinical decision making as it may help identify radioresistant versus radiosensitive tumors and influence the decision for definitive resection. Although single dimension tumor measurements using abdominal CT remains the primary criteria for assessment of treatment response, a number of studies have shown the synergistic role of PET to abdominal CT in response evaluation. As with other neoplasms, FDG uptake was found to significantly differentiate responders and non-responders and predict survival in patients with PA [8,20,21]. While the FDG uptake promisingly depicts tumor malignancy, rudimentary analyses, typically global $SUV_{max,}$ and SUV_{mean}, have limited accuracies in assessing the biologically complex PA tumors (Fig. 1). As demonstrated by genomic sequencing analysis, multiple tissue samples taken from the same tumor displayed a wide variety of cytogenetic differences [22,23]. Therefore, quantitative imaging methods describing the intra-tumor heterogeneity are needed.

In the past decade, fused PET/CT has become the standard imaging modality localizing FDG-intense foci. At the same time, more sophisticated PET/CT analysis algorithms have been developed to describe intra-tumor spatial heterogeneity. In particular, Tixier and colleagues calculated textural features, such as homogeneity and entropy, on baseline esophageal tumors and successfully predicted response to RT in 41 patients [24]. Yue and colleagues further compared pre- and post-RT intra-tumor PET textural features in 26 PA patients, and identified 3 borderline-significant prognostic variables, including homogeneity, variance and cluster tendency [25]. Instead of focusing only on FDG-avid regions, textural features describe the distribution of FDG uptake in the entire tumor region, which improved the characterization of tumor metabolic heterogeneity. However, in these PET textural studies, tumor volumes were only served as a binary mask for textural feature extraction, without contributing morphological information into the analyses. New alternatives that integrate both anatomic and metabolic information of tumor may provide a more complete picture of tumor response.

Tumors exhibit heterogeneity not only in space but also across time. In one hand, some PA cells grow faster than others, for instance, tumor cells in close proximity to pancreatic stellate cells have accelerated proliferation [26]. On the other, certain PA tissues, such as tumorigenic cancer cells defined by the CD133 expression, are more treatment-resistant than others [27]. Collectively, targeting the tumor microenvironment variations and probing into spatial-temporal coincidence in longitudinal imaging data may provide valuable insight of tumor treatment response [25,28]. In this study, we developed a 3D tumor surface based pipeline combing both the metabolic and morphological features to quantify tumor variations between pre- and post-RT. The tumor surface model characterizes the 3D tumor with regional spatio-temporal features. Comparing to univariate shape descriptors over the entire GTV, the use of 3D surface model facilitates comparisons of regional morphological features on locally aligned surface vertices, thus are sensitive to small shape alterations from tumor responses [13,14]. By projecting the SUV onto the tumor surface, regional metabolic features were integrated into the 3D model, and the intra-tumor metabolic heterogeneity was largely preserved in the response evaluation.

Our results showed that the combined morphological and metabolic metric, $M_{dist}|_{Fusion}$, significantly predicted OS and stratified risks for our PA patients, and outperformed traditional CA19-9, global SUV, and GTV based measurements. Furthermore, using the same 3D surface model, the morphological and metabolic fused metric $M_{dist}|_{Fusion}$ were superior to morphological or metabolic features alone ($M_{dist}|_{Shape}$, and $M_{dist}|_{SUV}$) in OS predicting and risk stratification. This, in turn, indicates the more accurate characterization of tumor response using the combined model. In addition to the PA patients, the presented pipeline may be applicable to other solid tumors for response evaluation.

There are several limitations in this study. First, in this pilot study,
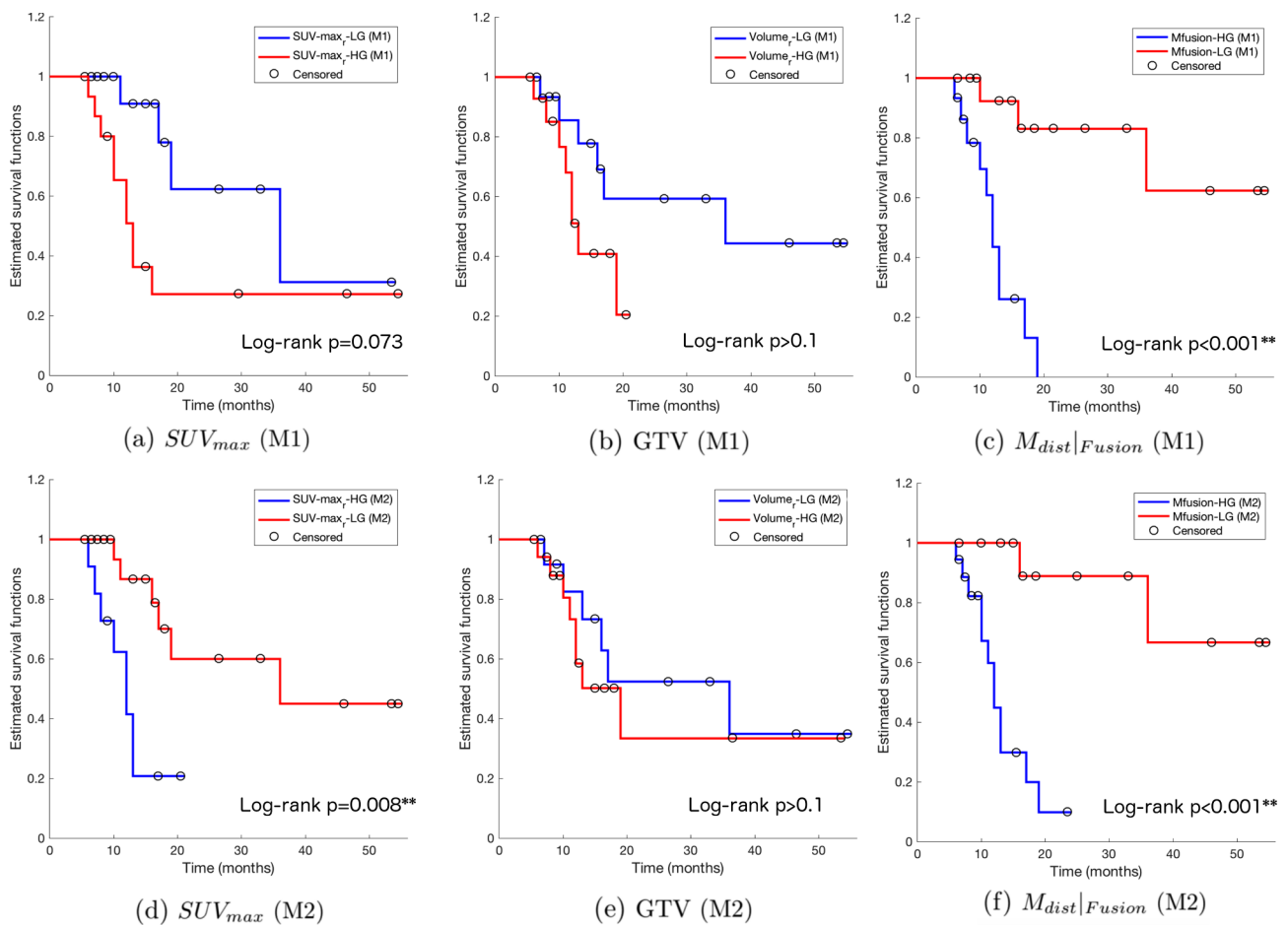
Fig. 3. Kaplan-Meier plots of OS for 31 patients stratified by post- versus pre-RT ratios of Global SUV_{max}, GTV, and the new $M_{dist}|_{Fusion}$, using M1, the evenly dividing method (a-c) and M2, the k-means clustering method (d-f). Censored observations are marked by black circles.

the enrolled participants were retrospectively solicited from our radiation oncology practice. The small sample size limits the power of statistical tests, and preclude the possibility to reliably remove the influence of all confounding factors. The comparisons between new multimodal based metric to traditional imaging metrics were made in the same condition and thus less biased by the heterogeneity of patient characteristics (tumor stage, radiotherapy fraction, censorship, etc.). However, the use of multiple PET/CT scanners may have led to some inconsistency in metabolic metric calculations despite our best effort to calibrate these scanners to be consistent. A calibration comparison

study on 11 regularly maintained PET/CT scanners observed good reproducibility and accuracy of all dose calibrators and reported no major deviation across different PET calibration systems [29]. Nevertheless, many quantitative descriptors draw from GTVs, for instance, textural features, have been reported to have limited robustness against multi-scanners [30]. Future larger cohorts preferably scanned in the same PET/CT scanner with high-resolution PET protocols, are needed to validate our method and findings. Additionally, segmentation accuracy is paramount to quantitative imaging analysis. A previous study investigating the reproducibility of PET radiomic features against GTV

Table 3
Results of log-rank tests for patient risk groups stratified by 6 imaging based metrics, using M1 and M2. Surface measurements based on metabolic features only, anatomic features only and combined anatomic and metabolic features are indicated as $M_{dist}|_{SUV}$, $M_{dist}|_{Shape}$, $M_{dist}|_{Fusion}$, respectively. Log-rank statistics of groups stratified using M1 and M2 are displayed in white and gray rows, respectively. Significance are marked as * for p-values < 0.05, and ** for p-values < 0.01.

	LG		HG		HR (95% CI)	p values
	N	Mean (Std)	N	Mean (Std)		
SUV_{max} (R)	16	0.51 (0.19)	15	1.12 (0.29)	3.11 (1.06–9.11)	0.073
	20	0.58 (0.22)	11	1.22 (0.27)	6.97 (1.94–25.07)	0.008**
SUV_{mean} (R)	16	0.68 (0.12)	15	1.16 (0.27)	2.78 (0.93–8.27)	> 0.1
	18	0.71 (0.14)	13	1.19 (0.27)	2.40 (0.80–7.24)	> 0.1
GTV (R)	16	0.63 (0.23)	15	1.19 (0.28)	2.78 (0.88–8.78)	> 0.1
	13	0.57 (0.22)	18	1.14 (0.28)	1.47 (0.50–4.34)	> 0.1
$M_{dist}|_{SUV}$	16	5.33 (6.13)	15	1.03 (0.44)	4.30 (1.42–13.02)	0.021*
	15	5.50 (6.32)	16	1.15 (0.54)	4.89 (1.56–15.29)	0.015*
$M_{dist}|_{Shape}$	16	1.18 (0.81)	15	0.17 (0.10)	1.90 (0.66–5.52)	> 0.1
	20	0.24 (0.15)	11	1.52 (0.76)	1.04 (0.35–3.08)	> 0.1
$M_{dist}|_{Fusion}$	16	4.65 (5.82)	15	0.57 (0.35)	11.51 (3.48–38.02)	< 0.001**
	13	5.37 (6.27)	18	0.73 (0.50)	9.04 (2.89–28.33)	< 0.001**

delineation has reported increased robustness of derived radiomic descriptors when using semi-automated segmentation over manual slice-by-slice delineation [31]. Same concerns also apply to the presented method, as it is sensitive to tumor volume segmentation errors that are prominent with less conspicuous tumors. With recent developments of deep learning, several automated tumor segmentation studies have shown promising results [32,33]. While accurate pancreatic tumor delineation leads to ongoing research, semi-automated or automated segmentation tools may be integrated into our pipeline to improve consistency and accuracy.

In conclusion, based on PET/CT, we introduced a novel combined morphologic and metabolic pipeline for treatment assessment of locally advanced or borderline resectable PA. The fused M_{dist} outperformed traditional morphologic, metabolic, and physiological measurements in OS prediction and risk stratification. The presented fused model shows the promise to be a new imaging-marker capable to improve characterization of tumor response and personalized PA care. Future large-scale studies will be needed to validate the capacity of the combined metric generated here.

Conflicts of interest

All authors confirm that there are no known conflicts of interest associated with this publication.

References

[1] Hidalgo M. Pancreatic cancer. N Engl J Med 2010;362(17):1605–17.
[2] Howlader N, Noone AM, Krapcho M, Miller D, Bishop K, Kosary CL, et al. SEER Cancer Statistics Review, 1975-2014, National Cancer Institute. Bethesda, MD, https://seer.cancer.gov/csr/1975_2014/, based on November 2016 SEER data submission, posted to the SEER web site, April 2017.
[3] Mian OY, Ram AN, Tuli R, Herman JM. Management options in locally advanced pancreatic cancer. Current Oncology Reports 2014;16(6):388.
[4] Katz MH, Fleming JB, Bhosale P, Varadhachary G, Lee JE, Wolff R, et al. Response of borderline resectable pancreatic cancer to neoadjuvant therapy is not reflected by radiographic indicators. Cancer 2012;118(23):5749–56.
[5] Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer 2009;45(2):228–247.
[6] Therasse P, Arbuck SG, Eisenhauer EA, Wanders J, Kaplan RS, Rubinstein L, et al. New guidelines to evaluate the response to treatment in solid tumors. J Natl Cancer Inst 2000;92(3):205–16.
[7] Cros J, Raffenne J, Couvelard A, Poté N. Tumor heterogeneity in pancreatic adenocarcinoma. Pathobiology 2018;85(1–2):58–65.
[8] Sahani DV, Bonaffini PA, Catalano OA, Guimaraes AR, Blake MA. State-of-the-art PET/CT of the pancreas: current role and emerging indications. Radiographics. 2012;32(4):1133–58.
[9] Jha P, Bijan B. PET/CT for pancreatic malignancy: potential and pitfalls. J Nuclear Med Technol 2015;43(2):92–7.
[10] Lemke AJ, Niehues SM, Hosten N, Amthauer H, Boehmig M, Stroszczynski C, et al. Retrospective digital image fusion of multidetector CT and 18F-FDG PET: clinical value in pancreatic lesions—a prospective study with 104 patients. J Nucl Med 2004;45(8):1279–86.
[11] Rabkin Z, Israel O, Keidar Z. Do hyperglycemia and diabetes affect the incidence of false-negative 18F-FDG PET/CT studies in patients evaluated for infection or inflammation and cancer? A comparative analysis. J Nucl Med 2010;51(7):1015–20.

[12] Ballehaninna UK, Chamberlain RS. The clinical utility of serum CA 19–9 in the diagnosis, prognosis and management of pancreatic adenocarcinoma: an evidence based appraisal. J Gastrointestinal Oncol 2012;3(2):105.
[13] Lorensen WE, Cline HE. Marching cubes: a high resolution 3D surface construction algorithm. ACM siggraph computer graphics 1987;21(4):163–9.
[14] Wang Y, Song Y, Rajagopalan P, An T, Liu K, Chou YY, et al. Alzheimer's Disease Neuroimaging Initiative. Surface-based TBM boosts power to detect disease effects on the brain: an N = 804 ADNI study. Neuroimage 2011;56(4):1993–2010.
[15] Wang Y, Zhang J, Gutman B, Chan TF, Becker JT, Aizenstein HJ, et al. Multivariate tensor-based morphometry on surfaces: application to mapping ventricular abnormalities in HIV/AIDS. NeuroImage 2010;49(3):2141–57.
[16] Lao Y, Wang Y, Shi J, Ceschin R, Nelson MD, Panigrahy A, et al. Thalamic alterations in preterm neonates and their relation to ventral striatum disturbances revealed by a combined shape and pose analysis. Brain Struct Funct 2016;221(1):487–506.
[17] Lao Y, David J, Mirhadi A, Lepore N, Sandler H, Wang Y, et al. Discriminating lung adenocarcinoma from lung squamous cell carcinoma using respiration-induced tumor shape changes. Phys Med Biol 2018;63(21):215027.
[18] Lao Y, Law M, Shi J, Gajawelli N, Haas L, Wang Y, Leporé A. A T1 and DTI fused 3D corpus callosum analysis in pre- vs. post-season contact sports players. In10th International Symposium on Medical Information Processing and Analysis 2015;9287:92870O.
[19] Lao Y, Nguyen B, Tsao S, Gajawelli N, Law M, Chui H, et al. A T1 and DTI fused 3D corpus callosum analysis in MCI subjects with high and low cardiovascular risk profile. NeuroImage: Clin 2017;14:298–307.
[20] Bang S, Chung HW, Park SW, Chung JB, Yun M, Lee JD, et al. The clinical usefulness of 18-fluorodeoxyglucose positron emission tomography in the differential diagnosis, staging, and response evaluation after concurrent chemoradiotherapy for pancreatic cancer. J Clin Gastroenterol 2006;40(10):923–9.
[21] Kuwatani M, Kawakami H, Eto K, Haba S, Shiga T, Tamaki N, et al. Modalities for evaluating chemotherapeutic efficacy and survival time in patients with advanced pancreatic cancer: comparison between FDG-PET, CT, and serum tumor markers. Intern Med 2009;48(11):867–75.
[22] Harada T, Okita K, Shiraishi K, Kusano N, Kondoh S, Sasaki K. Interglandular cytogenetic heterogeneity detected by comparative genomic hybridization in pancreatic cancer. Cancer Res 2002;62(3):835–9.
[23] Fisher R, Pusztai L, Swanton C. Cancer heterogeneity: implications for targeted therapeutics. Br J Cancer 2013;108(3):479.
[24] Tixier F, Le Rest CC, Hatt M, Albarghach NM, Pradier O, Metges JP, et al. Intratumor heterogeneity characterized by textural features on baseline 18F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer. J Nucl Med 2011;52(3):369.
[25] Yue Y, Osipov A, Fraass B, Sandler H, Zhang X, Nissen N, et al. Identifying prognostic intratumor heterogeneity using pre-and post-radiotherapy 18F-FDG PET images for pancreatic cancer patients. J Gastrointestinal Oncol 2017;8(1):127.
[26] Bachem MG, Zhou S, Buck K, Schneiderhan W, Siech M. Pancreatic stellate cells—role in pancreas cancer. Langenbeck's Arch Surg 2008;393(6):891–900.
[27] Hermann PC, Huber SL, Herrler T, Aicher A, Ellwart JW, Guba M, et al. Distinct populations of cancer stem cells determine tumor growth and metastatic activity in human pancreatic cancer. Cell Stem Cell 2007;1(3):313–23.
[28] Gillies RJ, Anderson AR, Gatenby RA, Morse DL. The biology underlying molecular imaging in oncology: from genome to anatome and back again. Clin Radiol 2010;65(7):517–21.
[29] Bouchet F, Geworski L, Knoop BO, Ferrer L, Barriolo-Riedinger A, Millardet C, et al. Calibration test of PET scanners in a multi-centre clinical trial on breast cancer therapy monitoring using 18F-FLT. PLoS ONE 2013;8(3):e58152.
[30] Traverso A, Wee L, Dekker A, Gillies R. Repeatability and reproducibility of radiomic features: a systematic review. Int J Radiation Oncol* Biol* Phys 2018;102(4):1143–58.
[31] Parmar C, Velazquez ER, Leijenaar R, Jermoumi M, Carvalho S, Mak RH, et al. Robust radiomics feature quantification using semiautomatic volumetric segmentation. PLoS ONE 2014;9(7):e102107.
[32] Gou S, Wu J, Liu F, Lee P, Rapacchi S, Hu P, et al. Feasibility of automated pancreas segmentation based on dynamic MRI. British J Radiol 2014;87(1044):20140248.
[33] Xing F, Xie Y, Yang L. An automatic learning-based framework for robust nucleus segmentation. IEEE Trans Med Imaging 2016;35(2):550–66.