



Preliminary exploration of response the course of radiotherapy for stage III non-small cell lung cancer based on longitudinal CT radiomics features

Ruiping Zhang^{a,1,2}, Zhengting Cai^{b,2,3}, Yan'an Luo^{c,4}, Zhizhen Wang^{a,*}, Wei Wang^{a,*}

^a Department of Radiation Oncology, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center of Cancer, Key Laboratory of Cancer Prevention and Therapy, Tianjin's Clinical Research Center for Cancer, Huanhu West Road, Hexi District, Tianjin 300060, China

^b School of Automation (Artificial Intelligence), Hangzhou Dianzi University, Xiasha Higher Education Zone, Hangzhou, Zhejiang Province 310018, China

^c Department of Physics, Nankai University, Weijin Road, Nankai District, Tianjin 300071, China

ARTICLE INFO

Keywords:

Non-small cell lung cancer
Longitudinal radiomics features
Radiation therapy
Computed tomography

ABSTRACT

Purpose: Explore the longitudinal CT-based radiomics to demonstrate the changing trend of radiotherapy response and to determine at which point after the onset of treatment radiomics exhibit the greatest change for stage III NSCLC patients.

Methods and materials: Ten stage III NSCLC patients in line with inclusion criteria were enrolled retrospectively, each of whom received radiotherapy or concurrent chemo-radiotherapy and performed eight series of follow-up CT imaging. Longitudinal radiomics were extracted on region of interest from the eight registered images, then two steps were conducted to select significant features as indicators of tumor change: 1) stable features were selected by Kendall rank correlation; 2) texture feature types with a steadily changing trend were retained and intensity features with stable change trends were selected to represent the large number of them. Next, the trend and rate of tumor change were analyzed using the Delta method and Curve-fitting method. Finally, the statistics in the distribution of stable features in patients were calculated.

Results: 675 stable features were selected from a total number of 1371 radiomics features, then 12 texture features types were retained and three intensity features were chosen to represent their own category. Among the final selected feature types, it was found that the two time points were weeks 1 and 3 with the higher rate of change. One patient had very few stable tumor features out of a total of 101 features, and the rate of change of features of another patient was conspicuously higher than the average level with number of 301 features.

Conclusion: The longitudinal CT radiomics could demonstrate the change trend of tumor and at which point exhibit the greatest change during radiotherapy, and potentially be used for treatment decisions concerning adaptive radiotherapy.

1. Introduction

Despite advances in prevention, detection, and treatment, lung

cancer remains a prevalent disease, and is associated with high mortality in both men and women worldwide [1]. Non-small cell lung carcinoma (NSCLC) is the most common subtype, counting for over 85% of lung

Abbreviations: NSCLC, Non-small cell lung carcinoma; CBCT, Cone-beam Computed Tomography; CT, Computed Tomography; VMAT, Volumetric Modulated Arc Therapy; GTV, Gross Tumor Volume; ROI, Region of Interest; HU, Hounsfield Units; IBEX, Imaging Biomarker Explorer; GLCM25/GLCM3, Gray Level Co-occurrence Matrix25/Gray Level Co-occurrence Matrix3; GLRLM25, Gray Level Run Length Matrix25; NID25/NID3, Neighborhood Intensity Difference25/Neighborhood Intensity Difference3; LASSO, Least Absolute Shrinkage and Selection Operator; PCA, Principle Component Analysis.

* Corresponding authors.

E-mail addresses: rzhang01@tmu.edu.cn (R. Zhang), jety2858@hdu.edu.cn (Z. Cai), luoya@nankai.edu.cn (Y. Luo), mdanderson_1@163.com (Z. Wang), weiwang_2@126.com (W. Wang).

¹ Postal address: Huanhu West Road, Hexi District, Tianjin 300060, China.

² (co-first authorship) contributed equally to this work. They worked together to select clinical data, analyze the results and solve problems encountered in this study and wrote the manuscripts.

³ Postal address: Xiasha Higher Education Zone, Hangzhou 310018, Zhejiang Province, China.

⁴ Postal address: Weijin Road, Nankai District, Tianjin 300071, China.

<https://doi.org/10.1016/j.ejro.2021.100391>

Received 7 September 2021; Received in revised form 28 October 2021; Accepted 9 December 2021

Available online 16 December 2021

2352-0477/© 2021 Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

cancers [1]. While patients with overall stage III NSCLC who are unable to undergo surgical resection may receive adjuvant chemotherapy and radiation therapy, assessment of early response during the course of treatment could potentially advise the application of response-adapted therapy in a timely manner. Such therapies may include an adaptive radiation therapy boost or individualized adjuvant chemotherapy according to their specific biological effects [2].

Quantitative characteristics of lung tumors based on medical imaging, namely radiomics features, have been widely explored and studied [3] and correlated with tumor histology, tumor stage and patient overall survival [4,5]. In order to better reflect the role of radiomics in assessing therapeutic effects, many researchers use the change in radiomics features at the time of treatment to predict prognoses for different types of tumor. Determination and assessment of feature change has been successful in predicting the prognosis of liver metastasis in colorectal cancer to chemotherapies [6], predicting the occurrence of radiation pneumonitis of patients with esophageal cancer [7] and predicting the overall survival of NSCLC patients [8]. However, the changes in radiomics features were used to predict the prognosis of tumor at few time points in such studies, even after at the end of therapy, whereas it is likely to miss the early modifications in the metabolic activity of tissue during the course of treatment. Unfortunately, patients often have varied tumor responses to radiation therapy due to differences in tumor type and other genetic and epigenetic factors [9]. The ability of using radiomics features to predict response may enable early termination of treatment in non-responding patients, thus preventing additional toxicity, allow for early changes in treatment [8,10] and further guide the clinical decision-making resulting in better outcomes [11], since a timely switch to an alternative treatment approach may lead to a better prognosis where radiotherapy is not effective.

Therefore, some researchers explored the utility of cone-beam computed tomography (CBCT), which is routinely used to verify the position of the patient before radiation delivery, to capture the changes in the tumor as early as possible. Carsten et al. analyzed the assessment of tumor regression on CBCT for NSCLC patients [12]. Furthermore, Van Timmeren et al., investigated that using radiomics features extracted from CBCT images acquired during radiation treatment was capable of reflecting the early treatment response [13]. Although CBCT images are normally used in the radiation therapy regimen, they have the drawback of inferior quality compared with conventional computed tomography (CT) images, making it more difficult to extract radiomics features more precisely.

To the best of our knowledge, no study has investigated the use of CT images to assess longitudinal changes in the tumor during the entire course of radiation therapy. To better identify tumor changes in a timely manner and not miss the hidden changes, the first aim of this work was to find at which point in time during treatment shows the maximum change of internal environment of tumor with the eventual goal of exploring the possibility of using longitudinal radiomics for early treatment adaptation. The second aim was to explore which radiomics features were the best predictive signatures of response to radiotherapy.

2. Materials and methods

2.1. Patients

We retrospectively reviewed the CT images and medical records for patients with stage III NSCLC who received radiation therapy and concurrent chemotherapy between May 2017 and October 2019. Review of data was done with the approval of the Tianjin Cancer Hospital Medical Ethics Committee, which allowed waiver of informed consent. Criteria for inclusion and exclusion were as follows:

Inclusion criteria: (a) existence of CT images from prior to treatment as well as during the course of radiotherapy, (b) definitive pathology of the stage III NSCLC, (c) receiving of radiotherapy with a prescribed dose of 66–74 Gy, delivered via volumetric modulated arc

therapy (VMAT) (Varian VitalBeam, 6 MV X-ray), (d) ages > 18 years, (e) Karnofsky performance status (KPS) \geq 70;

Exclusion criteria: (a) history of other lung comorbidities, (b) lung surgery before radiotherapy, (c) prior radiation therapy to the planned radiation therapy fields.

After screening, ten patients were included in the final data analysis. Clinical characteristics of these patients are shown in Table 1.

2.2. Image acquisition, parameter setting and sampling

The CT scanning parameters were set up as follows: peak tube voltage of 120kVp, tube current of 100 or 200 mA, and an exposure time of 500–800 ms on a large-bore CT scanner (Brilliant Philips Healthcare, Best, the Netherlands) system. Axial images were reconstructed in a 512 \times 512 matrix at an in-plane resolution of 0.98 mm and image thickness of 2.5 mm. Retrospective analysis of pre-therapy images (baseline, Week 0) and weekly CT images (Week 1, Week 2, Week 3, Week 4, Week 5, Week 6 and Week 7) was acquired at eight sampling time points. In total, each patient had eight sets of CT images for the study. All images were automatically resampled with cubic voxels of 3 \times 3 \times 3 mm³ with linear interpolation using an automatic method in a dedicated commercial software (MIM Software Inc, version 6.9.6).

2.3. Tumor segmentation

The gross tumor volume (GTV) was identified as the region of interest (ROI) in our study. The GTV was semi-automatically contoured on the clinical treatment planning system (Eclipse Varian Medical System, Palo Alto, CA) to avoid the inclusion of peripheral consolidation and air bubbles in the lesion. This approach has been validated as a robust way of delineation in our clinical practice. The GTV was also registered to each subsequent weekly CT scan using a deformable image registration procedure in the Eclipse system [14]. All contours were investigated to ensure consistency and modified to avoid the normal lung and bone structure as well as air bubbles using a lower threshold of – 100 Hounsfield Units (HU) and an upper threshold of 200 HU. Those patients whose ROI volume with less than 5 cm³ was excluded.

2.4. Feature extraction

Imaging Biomarker Explorer (IBEX) v1.0 β , an open infrastructure software platform developed by Zhang et al. [15] was used to extract radiomics features. Seven categories of features were identified: Intensity Direct (n = 55), Intensity Histogram (n = 49), Gray Level Co-occurrence Matrix25 (GLCM25, n = 330), Gray Level Co-occurrence Matrix3 (GLCM3, n = 924), Gray Level Run Length Matrix25 (GLRLM25, n = 33), Neighbor Intensity Difference25 (NID25, n = 5), and Neighborhood Intensity Difference3 (NID3, n = 5). Features GLCM25/3, GLRLM25 and NID25/3 are texture features with quantifying regional heterogeneity difference. Intensity Direct and Intensity Histogram are intensity features describing different intensities between voxels. The relevant parameters are defaulted in IBEX when extracting features of each category.

All of the following information is from the work of Zhang et al. [15]:

- (1) “25” means two and half dimensions, for which the specific individual intensity pairs are counted slice by slice in two directions and the final feature matrix is the summation of the intensity pairs, while “3” means three dimensions so those intensity pairs are counted in three directions.
- (2) Thirty features are repeated in the two categories of Intensity Direct and Intensity Histogram due to the identical calculation formula [16,17]; we only kept one of the two and meanwhile combined the two categories as a new one named Intensity (n = 74).

Table 1
Clinical characteristics of the ten patients in our study.

Patient ID	Gender	Age (yrs)	T stage	N stage	Smoking Status	Tumor Histology	KPS	Total Radiation Dose (Gy)
D1	M	≥ 65	T2b	N3	Yes	Squamous cell carcinoma	70–80	< 70
D2	M	≥ 65	T2b	N3	No	Squamous cell carcinoma	70–80	< 70
D3	F	< 65	T2a	N2	No	Adenocarcinoma or other	80–90	> 70
D4	F	≥ 65	T2b	N3	No	Adenocarcinoma or other	70–80	< 70
D5	F	≥ 65	T2B	N3	No	Adenocarcinoma or other	90–100	< 70
D6	M	< 65	T1b	N3	No	Squamous cell carcinoma	90–100	> 70
D7	M	< 65	T1b	N2	Yes	Squamous cell carcinoma	70–80	> 70
D8	M	< 65	T2b	N3	No	Squamous cell carcinoma	70–80	< 70
D9	M	< 65	T2b	N3	No	Adenocarcinoma or other	90–100	< 70
D10	F	≥ 65	T1b	N2	Yes	Adenocarcinoma or other	90–100	> 70

Note ID: identity number; F: female, M: male; yrs, years; KPS, Karnofsky performance status.

- (3) Twenty-two feature types can be calculated from one matrix of GLCM25 or GLCM3, and the distance and orientation are the key parameters to determine the feature metrics. For GLCM25, there are 3 distances (1, 4 and 7 units), and 4 orientations (0°, 45°, 90°, 135°) plus the average of all orientations for each distance (named as -333, same as below), so the number of GLCM25 features is $330 = 22 \times 3 \times (4 + 1)$; for GLCM3, there are also 3 distances (1, 4 and 7 units), but 12 orientations ($\varphi/\theta = 0^\circ/90^\circ, 90^\circ/90^\circ, 0^\circ/0^\circ, 45^\circ/90^\circ, 135^\circ/90^\circ, 90^\circ/45^\circ, 90^\circ/135^\circ, 0^\circ/45^\circ, 0^\circ/135^\circ, 45^\circ/54.7^\circ, 135^\circ/54.7^\circ, 45^\circ/125.3^\circ, 135^\circ/125.3^\circ$) plus the average of all orientations for each distance, so the number of GLCM3 features is: $924 = 22 \times 3 \times (13 + 1)$.
- (4) Eleven feature types are extracted from one matrix of GLRLM25, and the orientation is the vital parameter to determine the feature metrics. There are 0°, 90° and the average of two orientations, so the number of GLRLM25 features is: $33 = 11 \times 3$.

In short, a total of 1371 unique features [74 (Intensity) + 330 (GLCM25) + 924 (GLCM3) + 33 (GLRLM25) + 5 (NID25) + 5 (NID3)] were identified.

2.5. Feature selection

In this section, two main steps were conducted to select significant radiomics features as the useful indicators of longitudinal tumor change. The workflow of feature selection is shown in Fig. 1.

2.5.1. Select stable features

During the entire treatment, some features may demonstrate instability in a change trend. For example, the certain eigenvalue increasing during one interval and decreasing in another interval, unrepeatability among patients. Because this type of feature likely to be an impractical metric for evaluation of effect of radiotherapy, these features need to be removed from the feature cohort.

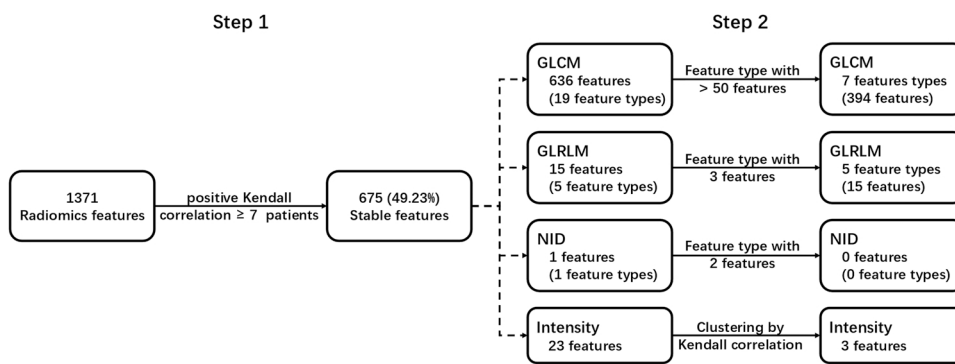


Fig. 1. The workflow of feature selection. Feature selection consisted of two steps, the first of which was to select the stable features with positive Kendall correlation in no fewer than 7 patients. Then the different feature categories were selected by different criteria and methods. Note: GLCM, GLRLM, and NID are texture features. GLCM (Gray Level Co-occurrence Matrix) includes the GLCM25 and GLCM3 with 22 feature types, in which each feature type has up to 57 features; GLRLM (Gray Level Run Length Matrix) is the GLRLM25 with 11 feature types, in which each feature type has up to 3 features; NID (Neighbor Intensity Difference) includes the NID25 and NID3 with 5 feature types, in which each feature type has up to 2 features.

(4) **Intensity:** Each intensity feature type only had one feature, so the above further selection method for texture features could no longer be utilized, so we clustered the features by Kendall rank correlation coefficients and selected the representative feature type in combination with the definition of features (Fig. 2).

2.6. Analysis

2.6.1. Analysis of changing trend and rate

In our study, the slope of the line obtained by linear regression was considered as the metric for the magnitude of feature change. We observed that the magnitudes of texture feature change for the same feature type were consistent and close even with different parameters of dimensions, distances and orientations (Fig. 3 and Table 2).

In order to acquire the changing rate over time and determine the time at which the changing rate reached a maximum, the following two methods were implemented:

(1) **Delta method**

For each selected feature type, the mean value of the feature value at each time point was calculated using its subordinate retained features, and differences between the mean values at time point t (from week 0 to week 6) and $t + 1$ were obtained. The

largest absolute difference corresponded to the time period when the feature change was most significant. The equation was:

$$|\Delta_{t+1}^i| = |\bar{x}_t^i - \bar{x}_{t+1}^i|$$

where i is one selected feature type, \bar{x} is the mean feature value of its subordinate retained features, t is the time point (from week 0 to week 6), Δ is the difference between the mean of two adjacent time points.

(2) **Curve-fitting method**

The change in each selected feature type was fitted by a cubic polynomial curve based on all of its retained features. Then, the derivative of the fitting curve was calculated, with the maximum absolute value corresponding to the time point with the fastest change.

2.6.2. Analysis of distribution of stable features in patients

Next, statistical properties in the distribution of stable features in patients were determined, including (1) the quantity of stable features in each patient, (2) the changing trends of stable features among patients, and (3) the detection of stable feature outliers. Analysis of the changing trends in stable features was accomplished by calculating the slopes of all stable features in ten patients regardless of whether a significant positive correlation was displayed or not, then comparing the trends (slope negative or positive) with each other and the mean level. The

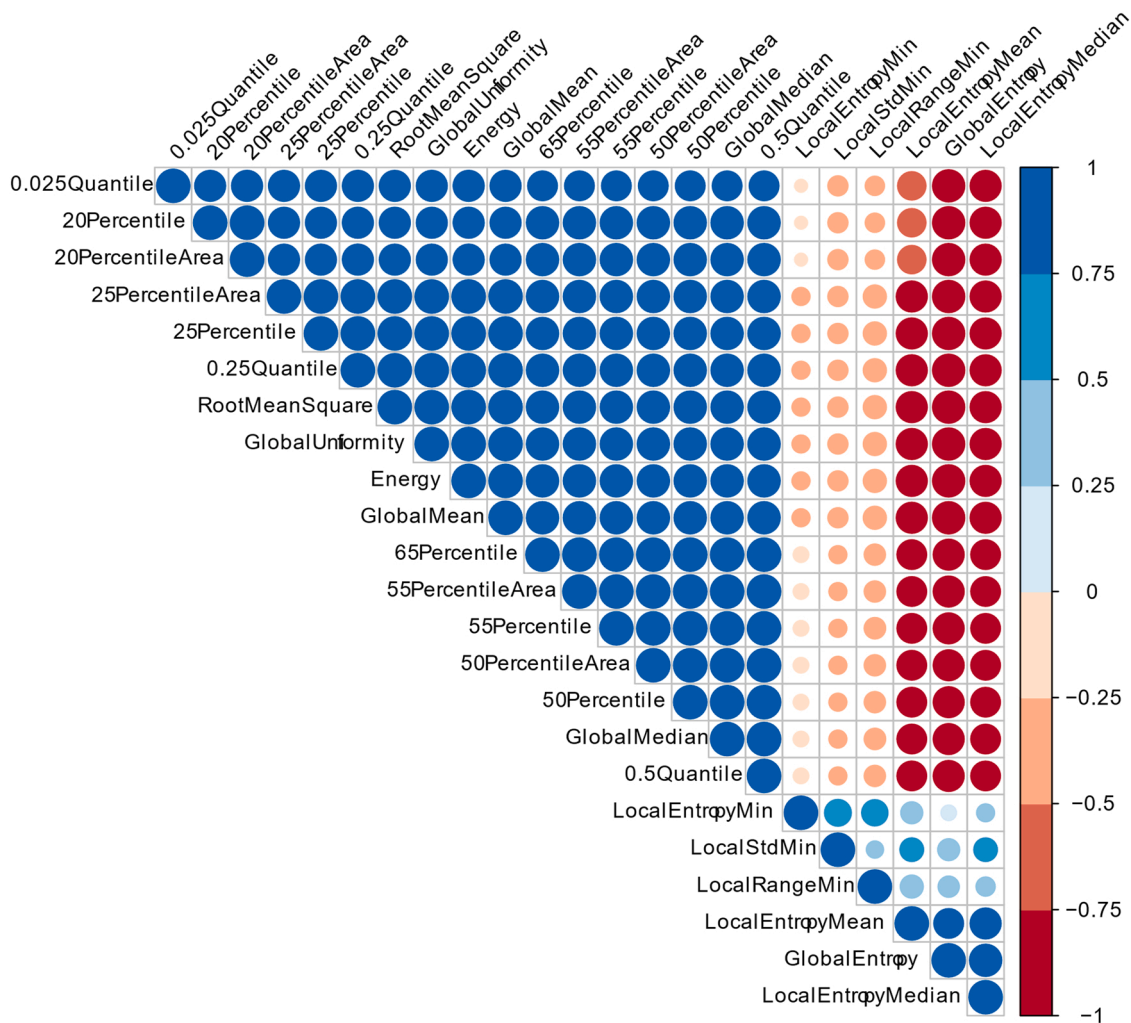


Fig. 2. The plot of correlation matrix of 23 stable intensity features. Circles with different colors and sizes represent the correlation coefficient between two features that are calculated based on the Kendall rank correlation. Then, the features are divided into three clusters according to the coefficients and a representative feature was selected in each cluster, which is 'Energy', 'GlobalEntropy', and 'LocalStdMin', respectively.

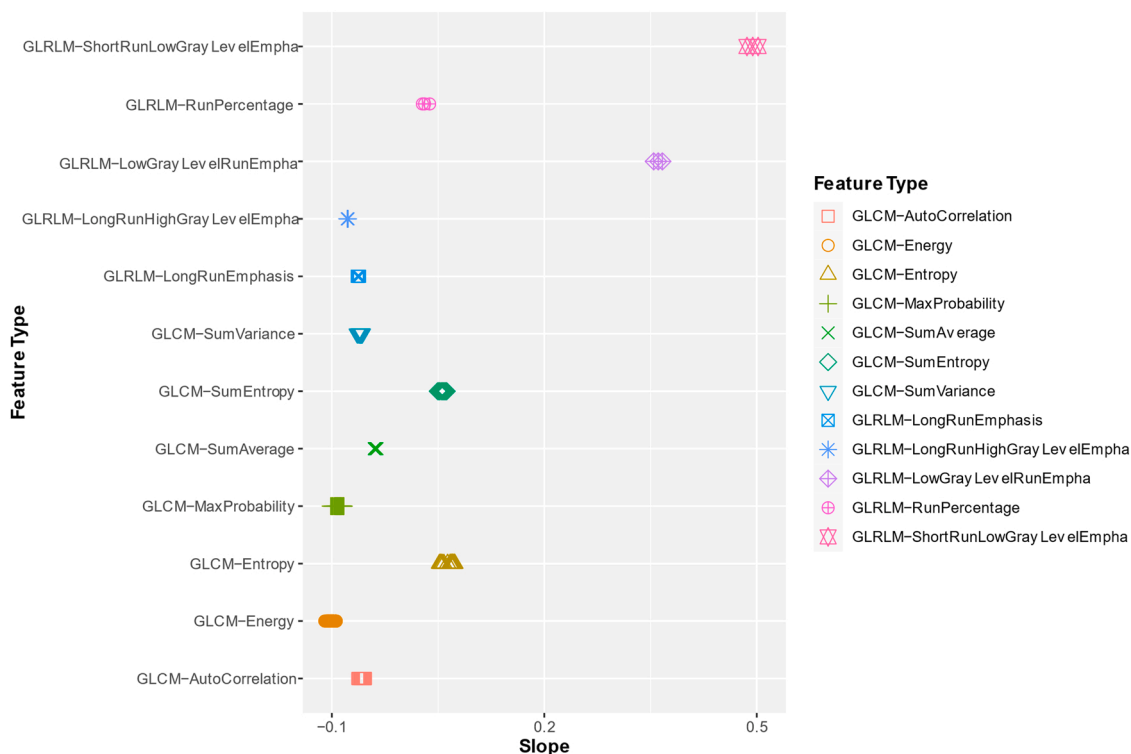


Fig. 3. The slopes of features in selected texture feature types. The slope of each feature was acquired by linear regression over time, considered as the changing trend of the feature. As we can see, in the same category and same feature type, the slopes of features are close (coefficient of variation for each type seen in Table 2), which means the changing trends of identical feature types are consistent and can be combined into a representation no matter what the parameters are. Note: GLCM (Gray Level Co-occurrence Matrix) and GLRLM (Gray Level Run Length Matrix) are the texture feature categories.

Table 2
The selected radiomics features (feature types) with their respective changing trend and rate.

Category	Feature/Feature type	Selection Ratio	Range of Slope	CV of Slope	Maximum changing period (Delta method)	Fastest change time point (Curve-fitting method)
GLCM	AutoCorrelation	57/57	-0.053 ~ -0.063	4.62%	Week 2 - Week 3	3.317 ∈ [Week 3 - Week 4]
	Entropy	57/57	0.052 ~ 0.073	4.30%	Week 2 - Week 3	0 ∈ [Week 0 - Week 1]
	Energy	57/57	-0.094 ~ -0.11	4.95%	Week 0 - Week 1	0 ∈ [Week 0 - Week 1]
	MaxProbability	57/57	-0.083 ~ -0.10	1.89%	Week 0 - Week 1	3.357 ∈ [Week 3 - Week 4]
	SumAverage	57/57	-0.036 ~ -0.040	2.60%	Week 2 - Week 3	3.256 ∈ [Week 3 - Week 4]
	SumEntropy	56/57	0.050 ~ 0.062	6.09%	Week 2 - Week 3	1.774 ∈ [Week 1 - Week 2]
	SumVariance	53/57	-0.057 ~ -0.064	2.60%	Week 2 - Week 3	0.429 ∈ [Week 0 - Week 1]
GLRLM	LongRunEmphasis	3/3	-0.061 ~ -0.064	2.47%	Week 0 - Week 1	0 ∈ [Week 0 - Week 1]
	LongRunHighGrayLevelEmpha	3/3	-0.077 ~ -0.078	0.80%	Week 0 - Week 1	0 ∈ [Week 0 - Week 1]
	LowGrayLevelRunEmpha	3/3	0.35 ~ 0.37	1.71%	Week 4 - Week 5	3.251 ∈ [Week 3 - Week 4]
	RunPercentage	3/3	0.029 ~ 0.038	16.20%	Week 0 - Week 1	0 ∈ [Week 0 - Week 1]
	ShortRunLowGrayLevelEmpha	3/3	0.48 ~ 0.50	1.58%	Week 4 - Week 5	3.321 ∈ [Week 3 - Week 4]
Intensity	Energy	1/1	-0.071	/	Week 2 - Week 3	3.152 ∈ [Week 3 - Week 4]
	GlobalEntropy	1/1	0.053	/	Week 2 - Week 3	1.604 ∈ [Week 1 - Week 2]
	LocalStdMin	1/1	0.13	/	Week 0 - Week 1	0 ∈ [Week 0 - Week 1]

Note: GLCM (Gray Level Co-occurrence Matrix) and GLRLM (Gray Level Run Length Matrix) are the texture feature categories, for which the names in the second column are the feature types, and the selection ratio of their subordinate features is represented by the form of ‘the number of features selected/the maximum number’. The slope of each feature in respective types was obtained by linear regression over time and considered as the changing trend of the feature, and CV is the coefficient of variation. Most of the fastest change points in time determined by the Curve-fitting method are not sampling time points, so the periods they belong to are also marked. For Intensity features, each type only has one feature, and the slash indicates none.

detection of stable features outliers was accomplished via a boxplot, (rstudio.com/).

All of the work was performed in the R language (version 3.6.1, <https://www.r-project.org/>) in RStudio (version 1.3, <http://www.rstudio.com/>).

3. Results

3.1. Feature selection

From a total of 1371 radiomics features, 675 (49.2%) stable features were retained based on the Kendall rank correlation criteria. Among these, 652 were texture features and 23 were intensity features (Fig. 1). Then, further features selection for different feature categories based on different criteria were determined. For texture features, seven feature types of GLCM, five feature types of GLRLM and none of NID were selected. For intensity features, three features including ‘Energy’, ‘GlobalEntropy’ and ‘LocalStdMin’ were chosen to represent this category. The details of selected features are summarized in Table 2 and the correlation matrix of 23 intensity features are shown in Fig. 2.

3.2. Analysis of changing trend and rate

Among fifteen final selected feature types, eight feature types show the upward trends over time while the other seven show decline. Specifically, ‘GLRLM-LowGrayLevelRunEmpha’ and ‘GLRLM-ShortRunLowGrayLevelEmpha’ are the two feature types with the most significant upward trends, with a range of slopes of 0.35–0.37, and 0.48–0.50, respectively. Also the features related to entropy (i.e., ‘GLCM-Entropy’, ‘GLCM-SumEntropy’ and ‘Intensity-GlobalEntropy’) all display a distinct rising trend. On the other hand, the feature types related to energy including ‘GLCM-Energy’ and ‘Intensity-Energy’ display downward trends (Table 2 and Supplement 1: Fig. S1).

The Curve-fitting method (Fig. 4A and Figs. S1–2 of Supplement 1) and the Delta method (Fig. 4B and Tables S1–1 of Supplement 1) were used to analyze the changing rates of selected radiomics features as a function of time and determine the time point or periods of maximum rate (Table 2). More details are shown in Fig. S2 of Supplement 1. According to Table 2, the maximum changing periods obtained by two methods intersect for eleven feature types, among which five are the same. Week 1 (8 times by the Curve-fitting method and 6 times by the Delta method) and Week 3 (6 times by the Curve-fitting method and 7

times by the Delta method) are the two time points with the highest frequency of occurrence. Differences and preconditions of the two methods, and the significant time points mentioned above will be expanded in the Discussion section.

3.3. Analysis of distribution of stable features in patients

In the table of numbers of stable features in each patient’s radiomics feature set (Supplement 2: Tables S2–1), each radiomics set from different patients (D7, D8, D9 and D10) has more than 600 stable features (675 in total), while the number of stable features for patients D2, D4 and D5 are far below the average, which are 95, 183 and 148, respectively.

According to Supplement 2: Tables S2–2, the number of features with consistent trends in all patients is 549, while the other 126 features have inconsistent trends in some patients. The features from four patients (D1, D3, D6, D7) all display consistent trends with the mean values, whereas the other six patients display features with more or less inconsistent trends, especially D5, with 101 such features. Notably, the features in certain patients always show inconsistency with the mean value as well as those of the majority of patients.

Statistics for outlier detection of changing trends (Supplement 2: Table S2–3) illustrate that the magnitude of the change in the patient D9 is significantly higher than that of the others (the number of outliers is 301) and far exceeds the mean level (585 over 1.5 times the mean). Few outliers were detected in other patients, but five patients (D2, D3, D4, D5, D7) have many features that were less than 0.5 times of the mean, with the number of 645, 586, 641, 598 and 336, respectively.

4. Discussion

Overall, our study achieved the established objectives, which were (1) to select robust longitudinal radiomics features to monitor the change of tumor characteristics during treatment, and (2) to investigate the trend and rate of changes to further determine the period of greatest change of tumor.

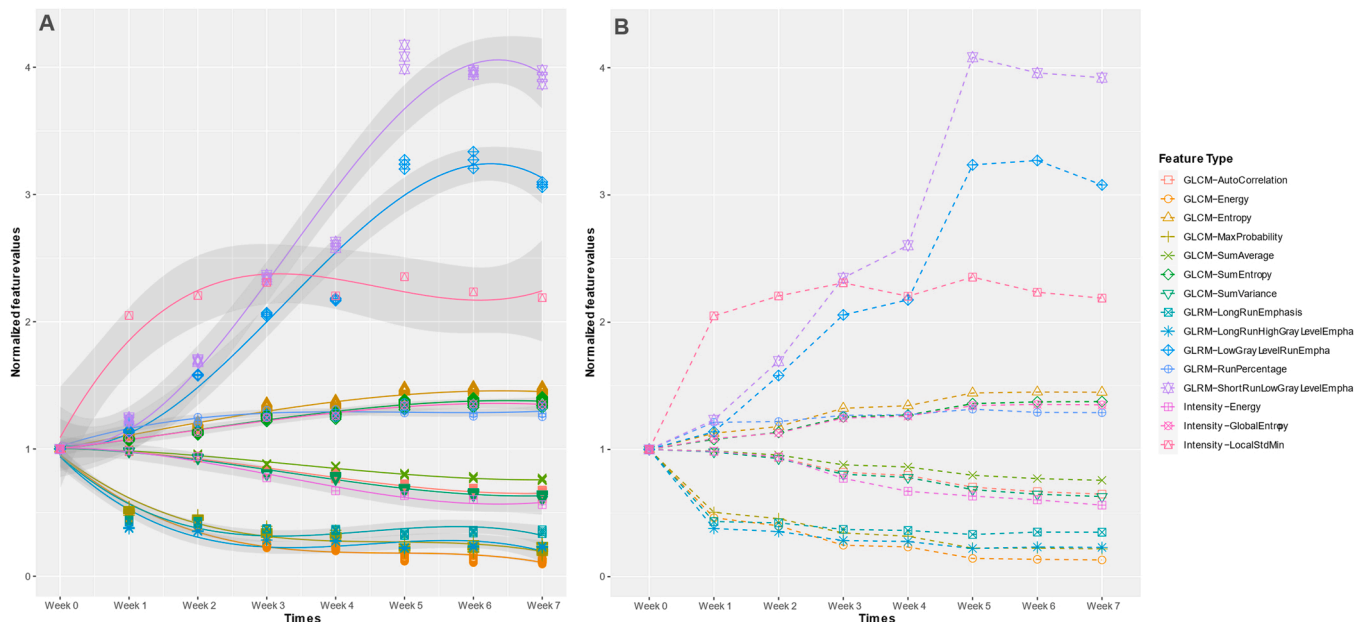


Fig. 4. The changes of selected radiomics features (feature types) as a function of time. In the subplot of A (Curve-fitting method), each point corresponds to the normalized feature value belonging to a certain feature type. The curves of feature types (solid lines) were fitted by the cubic polynomial with the standard error interval (gray shadow). In each curve, the largest absolute derivative is considered as the maximum changing rate. In the subplot of B (Delta method), each point corresponds to the mean of the normalized values of the features subordinate to a certain feature type. Every two adjacent points were connected by the dashed line, and the vertical difference is calculated. Among these differences, the largest absolute value corresponds to the most changing time period. Furthermore, each feature type is represented by a unique shape and color.

It should be noted that conventional feature selection methodologies, such as Least Absolute Shrinkage and Selection Operator (LASSO) and Principal Component Analysis (PCA) are somewhat inappropriate because of the very limited (ten patients), but valuable, data in our study. The method we describe in this paper presents features, not only based on single-point values, but also in the form of global values. Moreover, the method was easy to implement and effective for this study. First, considering the variability of some radiomics features owing to their intrinsic properties [3,18,19], the elimination of the unstable and unrepeatable features at the start has been part of the research paradigm in radiomics. In our study, the consistency of changing trends in the same cohort was utilized to discriminate whether the feature is stable or not. Then, we noticed that the texture features calculated by different parameters (such as dimension, angle, and displacement) from the same feature type displayed the approximate changing trends over time (seen in the coefficient of variation of slopes in Table 2 and scatter plot in Fig. 3), despite the different values at certain points in time. It just so happens that the significance of changing trend is far greater than the real feature value in the longitudinal analysis. Therefore, this phenomenon led us to view the data in terms of feature type rather than individual feature, thus simplifying and clarifying further feature selection and analysis.

The methods for calculating the changing rate and the maximum changing period include the Delta method and the Curve-fitting method. The former has been widely applied to the research on delta- or longitudinal radiomics [8,13,20], whereas the latter appears to have been rarely used and only by us up to date according to the published articles. From the results, the most significant change in features over time are mainly concentrated in two periods: one is at the very beginning of treatment, namely, from week 0 to week 1. This is in line with our previous expectation, because the tumor response to treatment between days 3–15 is indicated by decreased tumor growth accompanied with decreased vascular density [21]. The other is around the third week (Week 3), which is the highest frequency of occurrence for the maximum changing periods. By contrast, the change from Week 5 to Week 7 flattens out. The above findings draw forth an optimization direction of a radiotherapy or concurrent chemo-radiotherapy regimen for patients with stage III NSCLC towards precision and personalization. In addition to raising more concern over patients at treatment onset, it is also recommended that a systematic imaging examination be performed in the third week of treatment, so that the combination of multiple biomarkers including imaging-based radiomics can help accurately assess therapeutic response from patients and adjust the radiotherapy and treatment plan in time.

As a matter of fact, there are differences between the two methods by comparison. According to our practical experience, the conditions of use of the two methods are summarized as follows: The Delta method is more suitable for those feature types with a large amount of data. The strategies of fitting and averaging used in both the Curve-fitting and the Delta methods can effectively reduce the errors. However, the Curve-fitting method needs to combine the values at all time points for regression, whereas the Delta method considers the difference between each two adjacent sampling time points. Consequently, according to our analysis, the Delta method has the greater capability to reflect sharp variation of features in the entire data.

Taking the features of ‘GLCM-AutoCorrelation’ and ‘GLCM-Entropy’ as examples, these features change most significantly between Week 2 and Week 3, but after fitting, the maximum change time shifts forward or backward to some extent. On the other hand, the Curve-fitting method is more appropriate to those features with small numbers of data. In this case, the fitting method still retains the ability to reduce the effects of error, but averaging seems to help less. Besides, it is difficult to distinguish the internal variation in the data from errors in the data itself. ‘GLRLM-LowGrayLevelRunEmpha’ and ‘GLRLM-Short-RunLowGrayLevelEmpha’ are such examples, in which the data in Week 4 and 5 show apparent abnormalities and can be corrected by fitting. Of

course, under ideal conditions (with a sufficiently reliable patient database and a stable decision-making system), two or more different methods can be combined based on the idea of ensemble learning to make a comprehensive evaluation.

Furthermore, based on the distribution of stable features in patients, in three patients, D2, D4 and D5, only a few out of the 675 stable features were selected. This was especially true for patient D5, for whom the changing trend of 101 features showed inconsistency with the majority and the mean. Thus, we speculate that the original treatment might not have had a positive impact on the patients, being ineffective or even worse. Whereas the change trend for patient D9 was conspicuously greater than the average level, with 301 features detected as outliers by the boxplot method, we believe the effect of radiotherapy on this patient exceeded the normal expected effect. The relevant clinical reports roughly confirmed our speculations, but we cannot draw solid conclusions when lacking sufficient evidence and prior judgment in the retrospective study. We are hopeful that by responding to the above-mentioned appeal for adding imaging in the third week, the effect of treatment can be easily evaluated in the early and medium stage.

As far as we know, there are very few longitudinal studies in the field of radiomics, presumably due to the great difficulty of data acquisition; collecting the ten cases required a great deal of effort, and the only previous studies [8,13,20] mainly focused on the survival analysis based on CBCT imaging. More interestingly, the conclusions [20] about the performance of delta-radiomics to predict survival remain contradictory for the different processing methods. Even both pre- and post-reconstruction methods have been proposed to alleviate the effects of various artifacts on CBCT image quality [22–24], for example, scatter results in the occurrence of cup and streak artifacts [25]. Image artifacts may have other causes as well, such as a bad pixel in the KV imaging panel that may result in ring artifacts [26]. In addition, residual signals from previous acquisitions when visible on newly acquired images may result in ghost artifacts [27]. We admit that limitations of the present study lie in the quite small data size and the inability to standardize and obtain valid clinical information by retrospective analysis. Therefore, if possible, a prospective study should be carried out to acquire a larger imaging set and complete treatment records.

5. Conclusion

In this study, we have presented a preliminary attempt to evaluate tumor change trend during the course of radiotherapy based on longitudinal CT radiomics, and try to demonstrate at which point tumor exhibit the greatest change for stage III NSCLC. The method of analysis is efficient to explore the period of tumor maximum change and to select significant radiomics features for describing tumor reaction to treatment.

Funding

This work was supported by the Natural Science Foundation of Tianjin, China (20JCYBJC01510), and National Science Foundation of China, China (11875171, 81872472).

CRedit authorship contribution statement

Ruiping Zhang: Analyzing data, Writing – original draft, Writing – review & editing. **Zhengting Cai:** Methodology and Software. **Yanan Luo:** Data curation and Reviewing. **Wei Wang:** Visualization, Investigation and Reviewing. **Zhizhen Wang:** Conceptualization and Validation.

Acknowledgements

This work was supported by the Natural Science Foundation of

Tianjin, China (20JCYBJC01510), and National Science Foundation of China, China (11875171, 81872472).

The authors gratefully acknowledge Dr. George Starkschall from the Department of Radiation Physics, The University of Texas MD Anderson Cancer Center, Houston, TX77030, on the brilliant suggestion and revision of the manuscript.

Conflicts of interest

No conflict of interest to declare.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.ejro.2021.100391](https://doi.org/10.1016/j.ejro.2021.100391).

References

- [1] R.L. Siegel, K.D. Miller, A. Jemal, Cancer statistics, 2019, *CA Cancer J. Clin.* 69 (2019) 7–34.
- [2] T. Tirkes, M.A. Hollar, M. Tann, M.D. Kohli, F. Akisik, K. Sandrasegaran, Response criteria in oncologic imaging: review of traditional and new criteria, *Radiographics* 33 (2013) 1323–1341.
- [3] D.V. Fried, S.L. Tucker, S. Zhou, Z. Liao, O. Mawlawi, G. Ibbott, L.E. Court, Prognostic value and reproducibility of pretreatment CT texture features in stage III non-small cell lung cancer, *Int. J. Radiat. Oncol. Biol. Phys.* 90 (2014) 834–842.
- [4] H. Wang, X.H. Guo, Z.W. Jia, H.K. Li, Z.G. Liang, K.C. Li, Q. He, Multilevel binomial logistic prediction model for malignant pulmonary nodules based on texture features of CT image, *Eur. J. Radiol.* 74 (2010) 124–129.
- [5] B. Ganeshan, S. Abaleke, R.C. Young, C.R. Chatwin, K.A. Miles, Texture analysis of non-small cell lung cancer on unenhanced computed tomography: Initial evidence for a relationship with tumour glucose metabolism and stage, *Cancer Imaging* 10 (2010) 137–143.
- [6] S.X. Rao, D.M. Lambregts, R.S. Schnerr, R.C. Beckers, M. Maas, F. Albarello, R. G. Riedl, C.H. Dejong, M.H. Martens, L.A. Heijnen, W.H. Backes, G.L. Beets, M. S. Zeng, R.G. Beets-Tan, CT texture analysis in colorectal liver metastases: a better way than size and volume measurements to assess response to chemotherapy? *United Eur. Gastroenterol. J.* 4 (2016) 257–263.
- [7] A. Cunliffe, S.G. Armato, R. Castillo, N. Pham, T. Guerrero, H.A. Al-Hallaq, Lung texture in serial thoracic computed tomography scans: correlation of radiomics-based features with radiation therapy dose and radiation pneumonitis development, *Int. J. Radiat. Oncol. Biol. Phys.* 91 (2015) 1048–1056.
- [8] X. Fave, L. Zhang, J. Yang, D. Mackin, P. Balter, D. Gomez, D. Followill, A.K. Jones, F. Stingo, Z. Liao, R. Mohan, L. Court, Delta-radiomics features for the prediction of patient outcomes in non-small cell lung cancer, *Sci. Rep.* 7 (2017) 588, 584–11.
- [9] C.M. West, G.C. Barnett, Genetics and genomics of radiotherapy toxicity: towards prediction, *Genome Med.* 3 (2011) 52, 52–15.
- [10] T.H.L. Ralph, N. Georgi, C. Sara, et al., The effect of SUV discretization in quantitative FDG-PET Radiomics: the need for standardized methodology in tumor texture analysis, *Sci. Rep.* 5 (2015) 27.
- [11] P. Lambin, E. Rios-velazquez, R. Leijenaar, S. Carvalho, R.G. van Stiphout, P. Granton, C.M. Zegers, R. Gillies, R. Boellard, A. Dekker, H.J. Aerts, Radiomics: extracting more information from medical images using advanced feature analysis, *Eur. J. Cancer* 48 (2012) 441–446.
- [12] C. Brink, U. Bernchou, A. Bertelsen, O. Hansen, T. Schytte, S.M. Bentzen, Locoregional control of non-small cell lung cancer in relation to automated early assessment of tumor regression on cone beam computed tomography, *Int. J. Radiat. Oncol. Biol. Phys.* 89 (2014) 916–923.
- [13] J.E. Van Timmeren, R.T.H. Leijenaar, W. Van Elmpt, B. Reymen, P. Lambin, Feature selection methodology for longitudinal cone-beam CT radiomics, *Acta Oncol.* 56 (2017) 1537–1543.
- [14] H. Wang, L. Dong, M.F. Lii, A.L. Lee, R. de Crevoisier, R. Mohan, J.D. Cox, D. A. Kuban, R. Cheung, Implementation and validation of a three-dimensional deformable registration algorithm for targeted prostate cancer radiotherapy, *Int. J. Radiat. Oncol. Biol. Phys.* 61 (2005) 725–735.
- [15] L. Zhang, D.V. Fried, X.J. Fave, L.A. Hunter, J. Yang, L.E. Court, IBEX: an open infrastructure software platform to facilitate collaborative work in radiomics, *Med. Phys.* 42 (2015) 1341–1353.
- [16] H.J. Aerts, E.R. Velazquez, R.T. Leijenaar, C. Parmar, P. Grossmann, S. Carvalho, J. Bussink, R. Monshouwer, B. Haibe-Kains, D. Rietveld, F. Hoebers, M. M. Rietbergen, C.R. Leemans, A. Dekker, J. Quackenbush, R.J. Gillies, P. Lambin, Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach, *Nat. Commun.* 5 (2014) 4006–4009.
- [17] M. Ravanelli, D. Farina, M. Morassi, E. Roca, G. Cavalleri, G. Tassi, R. Maroldi, Texture analysis of advanced non-small cell lung cancer (NSCLC) on contrast-enhanced computed tomography: prediction of the response to the first-line chemotherapy, *Eur. Radiol.* 23 (2013) 3450–3455.
- [18] R.T.H. Leijenaar, S. Carvalho, E.R. Velazquez, W.J.C. van Elmpt, C. Parmar, O. S. Hoekstra, C.J. Hoekstra, R. Boellaard, A.L.A.J. Dekker, R.J. Gillies, H.J.W. L. Aerts, P. Lambin, Stability of FDG-PET radiomics features: an integrated analysis of test-retest and inter-observer variability, *Acta Oncol.* 52 (2013) 1391–1397.
- [19] X. Fave, D. Mackin, J. Yang, J. Zhang, D. Fried, P. Balter, D. Followill, D. Gomez, A. K. Jones, F. Stingo, J. Fontenot, L. Court, Can radiomics features be reproducibly measured from CBCT images for patients with non-small cell lung cancer? *Med. Phys.* 42 (2015) 6784–6797.
- [20] J.E. Van Timmeren, W. Van Elmpt, R. Leijenaar, B. Reymen, R. Monshouwer, J. Bussink, L. Paelinck, E. Bogaert, C. De Wagter, E. Elhaseen, Y. Lievens, O. Hansen, C. Brink, P. Lambin, Longitudinal radiomics of cone-beam CT images from non-small cell lung cancer patients: evaluation of the added prognostic value for overall survival and locoregional recurrence, *Radiother. Oncol.* 136 (2019) 78–85.
- [21] S.K. Kasoji, J.N. Rivera, R.C. Gessner, S.X. Chang, P.A. Dayton, Early assessment of tumor response to radiation therapy using high-resolution quantitative microvascular ultrasound imaging, *Theranostics* 8 (2018) 156–168.
- [22] M.C. Altunbas, C.C. Shaw, L. Chen, C. Lai, X. Liu, T. Han, T. Wang, A post-reconstruction method to correct cupping artifacts in cone beam breast computed tomography, *Med. Phys.* 34 (2007) 3109–3118.
- [23] Y. Zhang, L. Zhang, X.R. Zhu, A.K. Lee, M. Chambers, L. Dong, Reducing metal artifacts in cone-beam CT images by preprocessing projection data, *Int. J. Radiat. Oncol. Biol. Phys.* 67 (2007) 924–932.
- [24] S.F. Petit, W.J. Van Elmpt, S.M. Nijsten, P. Lambin, A.L. Dekker, Calibration of megavoltage cone-beam CT for radiotherapy dose calculations: correction of cupping artifacts and conversion of CT numbers to electron density, *Med. Phys.* 35 (2008) 849–865.
- [25] J.H. Siewerdsen, D.A. Jaffray, Cone-beam computed tomography with a flat-panel imager: magnitude and effects of x-ray scatter, *Med. Phys.* 28 (2001) 220–231.
- [26] R. Ning, X. Tang, D. Conover, R. Yu, Flat panel detector-based cone beam computed tomography with a circle-plus-two-arcs data acquisition orbit: preliminary phantom study, *Med. Phys.* 30 (2003) 1694–1705.
- [27] J.H. Siewerdsen, D.A. Jaffray, A ghost story: spatio-temporal response characteristics of an indirect-detection flat-panel imager, *Med. Phys.* 26 (1999) 1624–1641.