UC Irvine
UC Irvine Previously Published Works

Title

Increases in Serial Pretreatment 18F-FDG PET-CT Metrics Predict Survival in Early Stage Non-Small Cell Lung Cancer Treated With Stereotactic Ablative Radiation Therapy

Permalink

Journal

ISSN

Authors

Prionas, Nicolas D
von Eyben, Rie
Yi, Esther
et al.

Publication Date

DOI

Peer reviewed

Scientific Article

Increases in Serial Pretreatment ^{18}F-FDG PET-CT Metrics Predict Survival in Early Stage Non-Small Cell Lung Cancer Treated With Stereotactic Ablative Radiation Therapy

Nicolas D. Prionas MD, PhD [a,e], Rie von Eyben MSc [a], Esther Yi BA [a], Sonya Aggarwal BA [a], Jenny Shaffer MD [b], Jose Bazan MD [c], David Eastham MD [d], Peter G. Maxim PhD [a,e], Edward E. Graves PhD [a,e], Maximilian Diehn MD, PhD [a,e,f,*], Michael F. Gensheimer MD [a,e,*], Billy W. Loo Jr, MD, PhD [a,e,*]

[a]Department of Radiation Oncology, Stanford University School of Medicine, Stanford, California; [b]St. Anthony's Radiation Oncology Specialists, St. Anthony's Medical Center, St Louis, Missouri; [c]Department of Radiation Oncology, The James Cancer Hospital and Solove Research Institute, Columbus, Ohio; [d]David Grant Medical Center Radiation Oncology, Travis Air Force Base, Fairfield, California; [e]Stanford Cancer Institute, Stanford, California; and [f]Stanford Institute for Stem Cell Biology and Regenerative Medicine, Stanford, California

Abstract

Purpose: Quantitative changes in positron emission tomography with computed tomography imaging metrics over serial scans may be predictive biomarkers. We evaluated the relationship of pretreatment metabolic tumor growth rate (MTGR) and standardized uptake value velocity (SUVV) with disease recurrence or death in patients with early-stage non-small cell lung cancer treated with stereotactic ablative radiation therapy (SABR).

Methods and Materials: Under institutional review board approval, we retrospectively identified patients who underwent positron emission tomography with computed tomography at diagnosis and staging and simulation for SABR. Two cohorts underwent SABR between November 2005 to October 2012 (discovery) and January 2012 to April 2016 (validation). MTGR and SUVV were calculated as the daily change in metabolic tumor volume and maximum standardized uptake value, respectively. Cox proportional hazard models identified predictors of local, regional, and distant recurrence and death for the combined cohort. MTGR and SUVV thresholds dichotomizing risk of death in the discovery cohort were applied to the validation cohort.

Results: A total of 152 lesions were identified in 143 patients (92 lesions in 83 discovery cohort patients). In multivariable models, increasing MTGR trended toward increased hazard of distant recurrence (hazard ratio, 6.98; 95% confidence interval, 0.67-72.61; $P = .10$). In univariable models, SUVV trended toward risk of death (hazard ratio, 11.8, 95% confidence interval, 0.85-165.1, $P = .07$). MTGR greater than 0.04 mL/d was prognostic of decreased survival in discovery ($P = .048$) and validation cohorts ($P < .01$).
Conclusions: MTGR greater than 0.04 mL/d is prognostic of death in patients with non-small cell lung cancer treated with SABR. Increasing SUVV trends, nonsignificantly, toward increased risk of recurrence and death. MTGR and SUVV may be candidate imaging biomarkers to study in trials evaluating systemic therapy with SABR for patients at high risk of out-of-field recurrence.

Introduction

Treatment of medically inoperable early stage non-small cell lung cancer (NSCLC) with stereotactic ablative radiation therapy (SABR) results in 85% to 98% primary tumor control at 3 years.[1-3] There has been increasing evidence supporting the use of SABR in operable patients who have fewer comorbidities.[3,4] Although local recurrence is rare, the predominant pattern of failure is isolated distant recurrence[2] in 14% to 20% of cases at 5 to 7 years.[2,5]

[18]F-fluorodeoxyglucose positron emission tomography with computed tomography (^{18}F-FDG PET/CT) imaging is widely used for accurate staging[6] and has become increasingly prevalent for radiation therapy target delineation.[7] PET/CT for radiation therapy planning can provide new diagnostic information, including upstaging in 63% of locally advanced cases after 60 days.[8]

Prior studies have identified the predictive value of quantitative metrics on planning PET/CT. High pretreatment maximum standardized uptake value (SUV_{max}) and metabolic tumor volume (MTV) are predictive of disease progression and death.[9-11] Temporal changes in these parameters obtained from serial pretreatment PET/CT scans, expressed as SUV velocity (SUVV) and metabolic tumor growth rate (MTGR), might provide additional prognostic information as imaging correlates to tumor grade[12] and tumor doubling time,[13] respectively, which are known predictors of survival.[14-16] Because we routinely use PET/CT as part of simulation for SABR, in patients who typically previously underwent PET/CT during initial diagnosis and staging for lung cancer, these parameters may represent a readily available source of additional prognostic information in our patient population.

In this study we hypothesized that elevated SUVV and MTGR predict increased risk of recurrence or decreased survival. We primarily aimed to evaluate SUVV and MTGR in single predictor and multiple-predictor models of recurrence and survival. Secondarily, we aimed to identify threshold values of SUVV and MTGR that predict survival.

Materials and Methods

Patient population

Under institutional review board approval, we retrospectively identified adult patients with early stage biopsy-proven NSCLC treated with SABR in our department who underwent PET/CT imaging at 2 pretreatment time points: at diagnosis or staging and at radiation therapy simulation. All patients had no evidence of nodal disease and received SABR to the primary tumor. The discovery cohort was treated between November 2005 and October 2012 and the validation cohort was treated between January 2012 and April 2016. No exclusions were made based on age, gender, race, performance status, or length of time among staging, simulation, and treatment. Patient charts were retrospectively reviewed for clinical factors and clinical outcomes, including disease recurrence or death.

Treatment

All patients were treated with SABR, with definitive intent, and without chemotherapy. Treatment simulation was performed supine, with arms above the head and with custom immobilization. Simulation scans typically included both natural end-exhalation CT and 4-dimensional CT, as well as PET/CT. For inferior tumors, deep inspiratory breath hold was used. Gross tumor volume (GTV) was defined on thin-slice CT (1.25 mm slice thickness) using lung windowing. A clinical target volume expansion was not used. For tumors with minimal motion (<0.5-1 cm), a motion-inclusive internal target volume was defined. Mobile tumors were treated using a fluoroscopy guided amplitude-based motion management system (with or without implanted fiducial markers) or

image guided inspiratory breath-hold technique using previously described audiovisual biofeedback.[17,18] A 0.5-cm setup margin was added to define the planning treatment volume (PTV).

Patients in the discovery cohort were treated using CyberKnife (Acurray, Sunnyvale, CA), Varian Trilogy, or Varian TrueBeam (Varian, Palo Alto, CA) treatment systems. Patients in the validation cohort were treated using a Varian TrueBeam system only. SABR at our institution consists of delivering 25 to 60 Gy in 1 to 8 fractions, using a volumetric modulated arc therapy technique with 6 to 10 MV photons. Plans were normalized such that 95% of the PTV was covered by the prescription dose. Daily treatment setup included orthogonal planar kilovolt imaging and on-board cone beam CT with repositioning before treatment delivery.

^{18}F-FDG PET/CT imaging

PET/CT imaging was performed at diagnosis or staging and at treatment simulation. No restrictions were placed on the PET/CT system used. Staging imaging was performed on GE Discovery LS, GE Discovery ST, GE Discovery STE (GE Medical Systems, Milwaukee, WI), Philips Gemini TrueFlight (Philips Healthcare, Andover, MA), Siemens BioGraph HiRes Model 1080, and Siemens Biograph TruePoint Model 1093 (Siemens, Erlangen, Germany) scanners. Before 2012, simulation scans were performed on a GE Discovery ST platform, and after 2012 a Siemens Somatom Definition AS platform was used. The PET/CT protocol has been previously described[11] and consists of 8-hour fasting with blood glucose less than 160 mg/dL before injection of 10 to 18 mCi FDG, 45 to 60 minutes of tracer uptake time, and PET/CT acquisition with helical CT for attenuation correction.

PET/CT image analysis

All PET/CT images were evaluated using MiM Version 6.6.5 (Cleveland, OH) to measure SUV_{max} and MTV at diagnosis/staging (SUV_{Dx}, MTV_{Dx}) and at simulation (SUV_{Sim}, MTV_{Sim}); MTV was measured using the PET Edge (MiM Version 6.6.5) gradient-based segmentation tool consistently by a single observer. This technique defines the tumor edge by the largest signal gradient and corresponds with pathologic specimens more closely than threshold-based techniques.[19] Necrotic regions were included in the MTV. The change in SUV_{max} and MTV per unit time in days was calculated to define the SUVV and MTGR, respectively. Local recurrence (LR) was defined as biopsy-proven recurrence within the PTV or increase in the size and focal FDG-avidity within the PTV leading to a change in clinical management. Regional recurrence (RR) was defined as biopsy-proven

recurrence outside the PTV but involving the same lobe, ipsilateral hilum, mediastinum or supraclavicular region. Distant recurrence (DR) was defined as biopsy-proven recurrence in any region beyond local or regional recurrence.

Statistical analysis

The entire cohort of patients was used to develop univariable and multivariable Cox proportional hazard models of LR, RR, DR, and overall survival (OS). Benjamini-Hochberg false discovery rate adjustment was used. Multivariable Cox proportional hazard models included SUVV, MTGR, patient age, and MTV_{Sim} to test the independent predictive value of temporally inclusive metrics in the setting of known predictors (e.g., MTV). Patient, lesion, and imaging characteristics between the discovery and validation cohorts were also compared using Student's t test. Threshold values of SUVV and MTGR that maximize OS difference were identified in the discovery cohort using X-tile graphical software.[20] Validation of thresholds was attempted on the validation cohort. All statistics were performed in SAS Version 9.4 (SAS Institute Inc, Cary, NC).

Results

Patient, lesion, and treatment characteristics

A total of 152 lesions were evaluated in 143 patients: 92 lesions in 83 discovery cohort patients and 60 lesions in 60 validation cohort patients (Table 1). Patient and lesion characteristics were similar between cohorts, although GTV was significantly larger in the validation cohort ($P = .02$). The median biologically effective dose delivered was similar in both cohorts and greater than 100 Gy. The median prescription dose and fractionation was 50 Gy delivered in 4 fractions. Of note, a common SABR dose used to treat small peripheral tumors in our cohort is 25 Gy in a single fraction, which has a biologically effective dose of 87.5 Gy ($\alpha/\beta = 10$).

Of the discovery cohort, 22 of 83 (26.5%, 15.4% of full cohort) patients were treated on a CyberKnife platform. Lesions treated by CyberKnife were smaller than those treated on C-arm linear accelerator platforms with mean GTV of 10.6 mL (standard deviation, 11.0) versus 15.8 mL (standard deviation, 19.5), respectively, and with 15 of 22 (68.2%) lesions versus 76 of 130 (58.5%) lesions having volume smaller than 12 mL.

PET/CT image analysis

Sample patient images show cases in which rapidly changing pretreatment SUVV and MTGR (Fig. 1)

Table 1   Patient, lesion, treatment, and imaging characteristics

Characteristic	Discovery cohort	Validation cohort	All patients
	Median (range) or n (%)	Median (range) or n (%)	Median (range) or n (%)
Patient/lesion			
No. of patients	83	60	143
No. of lesions	92	60	152
Age at treatment end (y)	76 (42-99)	75 (47-91)	76 (42-99)
Female	44 (53)	28 (47)	72 (50)
KPS	80 (50-100)	80 (50-100)	80 (50-100)
Histologic type			
Adenocarcinoma	50 (54)	41 (68)	90 (60)
Squamous cell	29 (32)	11 (18)	40 (26)
Other	13 (14)	8 (14)	21 (14)
T stage			
T1	64 (70)	38 (63)	102 (67)
T2	26 (28)	22 (37)	48 (32)
T3	2 (2)	0 (0)	2 (1)
Treatment			
Diagnosis to treatment (d, interquartile range)	80 (50-119)	63 (39-82)	70 (44-108)
Total dose (Gy)	50 (25-60)	50 (25-60)	50 (25-60)
Fractions	4 (1-5)	3 (1-8)	4 (1-8)
BED	113 (80-180)	105 (80-151)	109 (80-180)
GTV (mL)*	8.6 (0.3-79.9)	11.5 (0.7-150.5)	9.4 (0.3-150.5)
PET/CT imaging			
Days between scans (interquartile range)	62 (34-99)	53 (28-69)	56 (30-86)
SUV (U)			
Diagnosis/staging	7.3 (0.9-23.4)	6.2 (1.4-22.4)	7.1 (0.9-23.4)
Simulation	9.0 (0.7-33.1)	6.8 (1.3-24.2)	8.2 (0.7-33.1)
MTV (mL)			
Diagnosis/staging	3.0 (0.1-39.0)	4.4 (0.3-57.3)	3.7 (0.1-57.3)
Simulation	5.1 (0.02-106.7)	7.9 (0.2-125.7)	5.9 (0.02-125.7)
SUVV (U/d)*	0.02 (−0.44-0.34)	0.01 (−1.28-0.32)	0.02 (−1.28-0.34)
MTGR (mL/d)	0.01 (−0.44-0.70)	0.02 (−0.31-0.88)	0.01 (−0.44-0.88)

Abbreviations: BED = biologically effective dose; GTV = gross tumor volume; KPS = Karnofsky performance status; MTGR = metabolic tumor growth rate; MTV = metabolic tumor volume; PET/CT = positron emission tomography with computed tomography; SUV = standardized uptake value; SUVV = SUV velocity.

  * *P* < .05 between discovery and validation cohorts.

resulted in distant recurrence (Fig. 1a, b) and slowly changing metrics did not recur after SABR (Fig. 1c, d).

For the entire cohort, the median duration between scans was 56 days (interquartile range 30-86). SUV$_{max}$ and MTV were not significantly different between cohorts, and the median value of each metric increased from diagnosis and staging to simulation (Table 1). For the full cohort, median (range) SUVV and MTGR were 0.02 U/d (−1.28 to 0.34) and 0.01 mL/d (−0.44 to 0.88), respectively. SUVV was significantly higher in the discovery cohort compared with the validation cohort (*P* < .05).

Clinical outcomes

Median follow-up was 30 months (range, 0-79 months). Median OS for the entire cohort was 47 months, and 2-year OS was 76.3%. The cumulative incidences of LR, RR, and DR at 2 years were 13.2%, 21.7%, and 19.0% and at 5 years were 25.6%, 36.3%, and 29.5%, respectively. For the discovery and validation cohorts, median follow-up was 36 and 16 months, respectively. Kaplan-Meier OS was not different between groups (*P* = .73). Median OS was 47 months for the discovery cohort and was not reached for the validation cohort; 2-year OS was 80.1% and 68.8% and 3-year OS was 67.4% and 62.4%, for the discovery and validation cohorts, respectively (Fig. 2a). The cumulative incidences of LR, RR, and DR at 2 years were 20.8%, 36.8%, and 32.7% for the discovery cohort and 10.7%, 14.8%, and 16.2% for the validation cohort, respectively. The cumulative incidences of LR, RR, and DR were significantly higher in the discovery cohort (*P* ≤ .01; Fig. 2).

Univariable and multivariable Cox models

In univariable Cox proportional hazard models on the combined cohort, increasing SUVV trended toward a
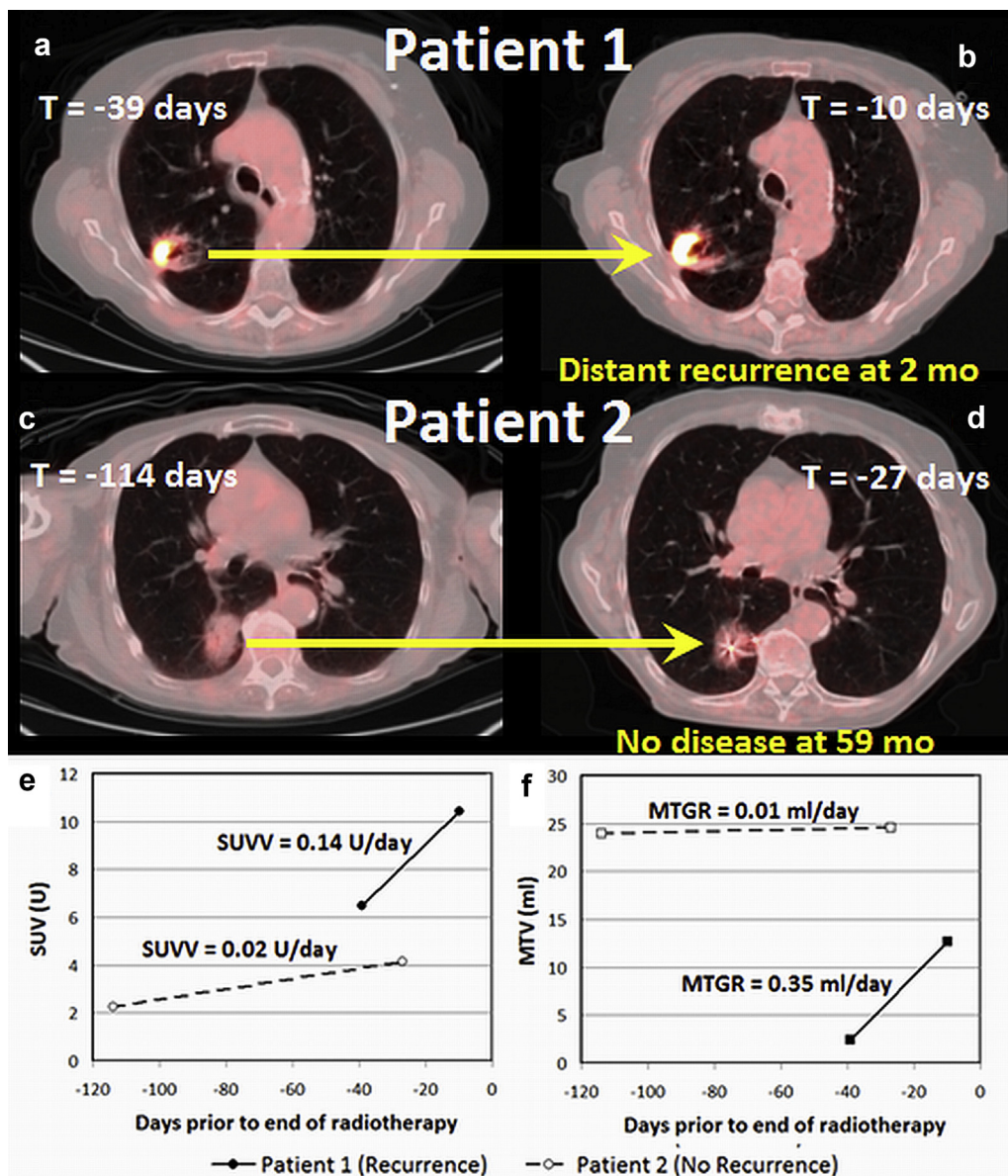
Fig. 1 Diagnostic and staging (a) and simulation (b) positron emission tomography with computed tomography (PET/CT) scans of an 81-year-old woman with T2 N0 adenocarcinoma of the right upper lung (patient 1) indicate high standardized uptake value velocity (SUVV) and metabolic tumor growth rate (MTGR; 0.14 U/d and 0.35 mL/d, respectively); she underwent stereotactic ablative radiation therapy (SABR; 50 Gy in 4 fractions) and developed distant metastases in the contralateral lung 2 months after treatment, regional recurrence in the right lower lung 14 months after treatment, and ultimately diffuse metastases. Conversely, the diagnostic and staging PET/CT scans (c) and simulation PET/CT scans after fiducial placement (d) in a 99-year-old man with T2 N0 adenocarcinoma of the right lower lung (patient 2) indicate low SUVV and MTGR (0.02 U/d and 0.01 mL/d, respectively); he underwent SABR (50 Gy in 4 fractions) and remained free from recurrence after 59 months of follow-up. Standardized uptake value (SUV) and metabolic tumor volume (MTV) at each time point, as measured from Fig. 1a, b, c, d, are shown in (e) and (f), respectively.

significant increase in hazard of death (hazard ratio [HR], 11.8; 95% confidence interval [CI], 0.85-165.1; $P = .07$). SUVV and MTGR were not significantly associated with hazard of LR, RR, or DR. Increasing GTV (HR, 0.98; 95% CI, 0.96-1.00; $P = .03$), MTV_{Dx} (HR, 0.95; 95% CI, 0.90-0.99; $P = .02$), and MTV_{Sim} (HR, 0.97; 95% CI, 0.95-1.00; $P = .03$) were associated with decreased

hazard of DR. After Benjamini-Hochberg false discovery rate adjustment, no univariable models were statistically significant, although the association between GTV, MTV, and SUVV and DR trended toward significance ($P = .11$; Table 2).

In multivariable Cox proportional hazard models, there was a trend toward increased hazard of DR for MTGR
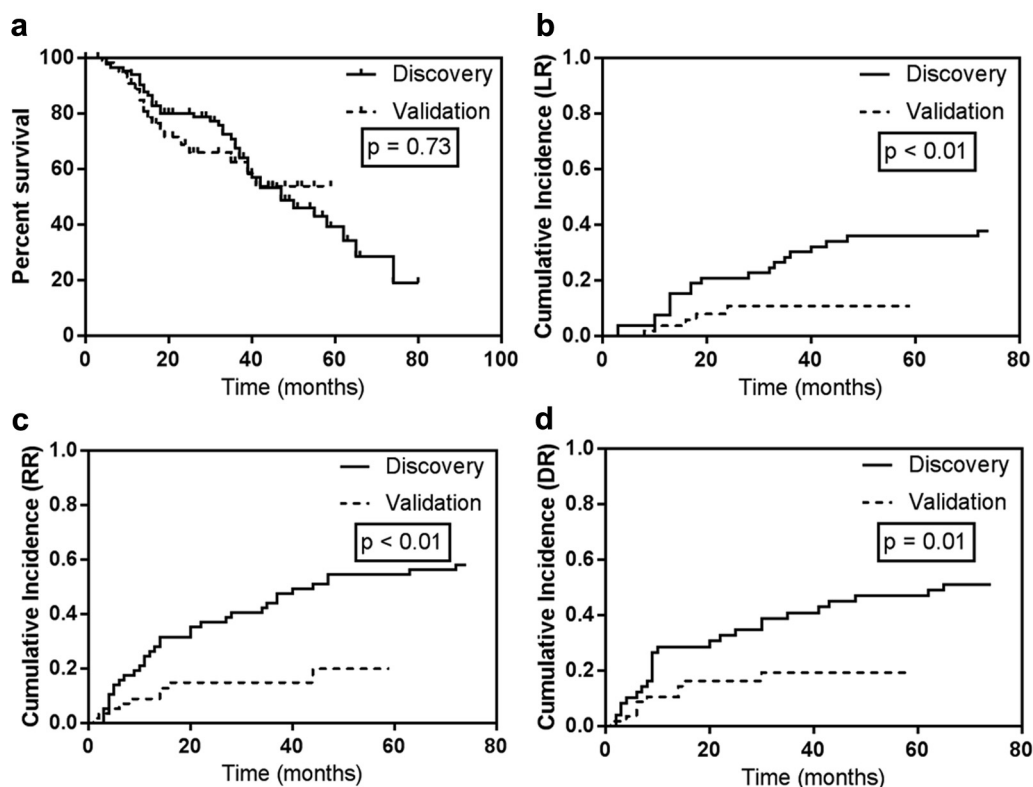
Fig. 2   Kaplan-Meier analysis of overall survival (a) and cumulative incidences of local recurrence (LR; b), regional recurrence (RR; c), and distant recurrence (DR; d) for the discovery and validation cohorts.

(HR, 6.98; 95% CI, 0.67-72.61; $P = .10$). Increasing SUVV was found to have a large but statistically insignificant hazard for LR ($P = .16$), RR ($P = .17$), DR ($P = .17$), and OS ($P = .12$). MTV remained associated with decreased hazard of DR (HR, 0.96; 95% CI, 0.93-0.99; $P = .01$; Table 3).

SUVV threshold for survival

Survival was not associated with SUVV or MTGR as continuous variables, so threshold values of SUVV and MTGR, to maximally dichotomize high versus low risk of death, were identified in the discovery cohort. SUVV less

Table 2   Univariable Cox proportional hazard regression models with Benjamini-Hochberg false discovery rate (FDR) adjustment

	Local recurrence		Regional recurrence		Distant recurrence			Overall survival	
	HR (95% CI)	*P*	HR (95% CI)	*P*	HR (95% CI)	*P*	FDR	HR (95% CI)	*P*
Clinical factors									
Age	0.99 (0.95-1.04)	.77	0.97 (0.93-1.01)	.12	0.99 (0.95-1.03)	.56	0.81	1.03 (0.99-1.06)	.10
T stage	0.55 (0.20-1.49)	.24	1.09 (0.66-1.78)	.74	0.78 (0.37-1.61)	.49	0.81	1.09 (0.63-1.90)	.75
Treatment factors									
BED	1.00 (0.99-1.01)	.98	1.00 (0.99-1.01)	.89	1.00 (0.99-1.01)	.99	0.99	1.00 (0.99-1.01)	.51
GTV	0.99 (0.97-1.02)	.53	0.99 (0.97-1.01)	.35	0.98 (0.96-1.00)	**.03**	0.11	1.00 (0.98-1.01)	.51
Imaging metrics									
SUV$_{Dx}$	1.04 (0.98-1.11)	.20	1.04 (0.99-1.09)	.15	1.02 (0.96-1.08)	.52	0.81	1.01 (0.97-1.06)	.61
SUV$_{Sim}$	1.04 (0.99-1.10)	.15	1.05 (1.00-1.10)	.06	1.01 (0.96-1.08)	.65	0.65	1.03 (0.99-1.08)	.12
MTV$_{Dx}$	0.99 (0.95-1.04)	.69	0.99 (0.95-1.02)	.43	0.95 (0.90-0.99)	**.02**	0.11	0.99 (0.97-1.02)	.62
MTV$_{Sim}$	0.99 (0.96-1.02)	.47	0.99 (0.97-1.01)	.27	0.97 (0.95-1.00)	**.03**	0.11	1.00 (0.99-1.01)	.61
SUVV	12.36 (0.38-404.55)	.16	4.34 (0.44-42.42)	.21	4.00 (0.38-42.57)	.25	0.63	11.8 (0.85-165.1)	.07
MTGR	0.63 (0.13-2.99)	.56	0.51 (0.14-1.93)	.32	0.78 (0.18-3.33)	.73	0.73	1.07 (0.29-3.94)	.91

Abbreviations: BED = biologically effective dose; CI = confidence interval; Dx = at diagnosis; GTV = gross tumor volume; HR = hazard ratio; MTGR = metabolic tumor growth rate; MTV = metabolic tumor volume; PET/CT = positron emission tomography with computed tomography; Sim = at simulation; SUV = standardized uptake value; SUVV = SUV velocity.
P-values less than 0.05 are bolded.

Table 3 Multivariable Cox proportional hazard regression models

Variable	Local recurrence		Regional recurrence		Distant recurrence		Overall survival	
	HR (95% CI)	P	HR (95% CI)	P	HR (95% CI)	P	HR (95% CI)	P
SUVV	12.03 (0.36-402.74)	.16	5.43 (0.58-62.03)	.17	5.98 (0.47-76.47)	.17	8.30 (0.60-115.59)	.12
MTGR	0.67 (0.13-3.31)	.62	0.73 (0.09-5.80)	.76	6.98 (0.67-72.61)	.10	2.78 (0.36-21.16)	.32
Age			0.96 (0.92-1.01)	.09	0.99 (0.95-1.04)	.68	1.03 (0.99-1.06)	.18
MTV$_{Sim}$			0.99 (0.97-1.02)	.51	0.96 (0.93-0.99)	**.01**	0.99 (0.97-1.01)	.32

Abbreviations: CI = confidence interval; HR = hazard ratio; MTGR = metabolic tumor growth rate; MTV$_{Sim}$ = metabolic tumor volume at simulation; SUVV = standardized uptake value velocity.
P-values less than 0.05 are bolded.

than 0.08 U/d was associated with significantly longer median OS (57 vs 32 months, $P < .01$; Fig. 3a) and MTGR less than 0.04 mL/d had significantly longer median OS (57 vs 38 months, $P = .048$; Fig. 3b). When applied to the validation cohort, the SUVV threshold was not significantly associated with survival ($P = .43$; Fig. 3c); MTGR threshold was significantly associated with a survival difference ($P < .01$; Fig. 3d). Of note, the
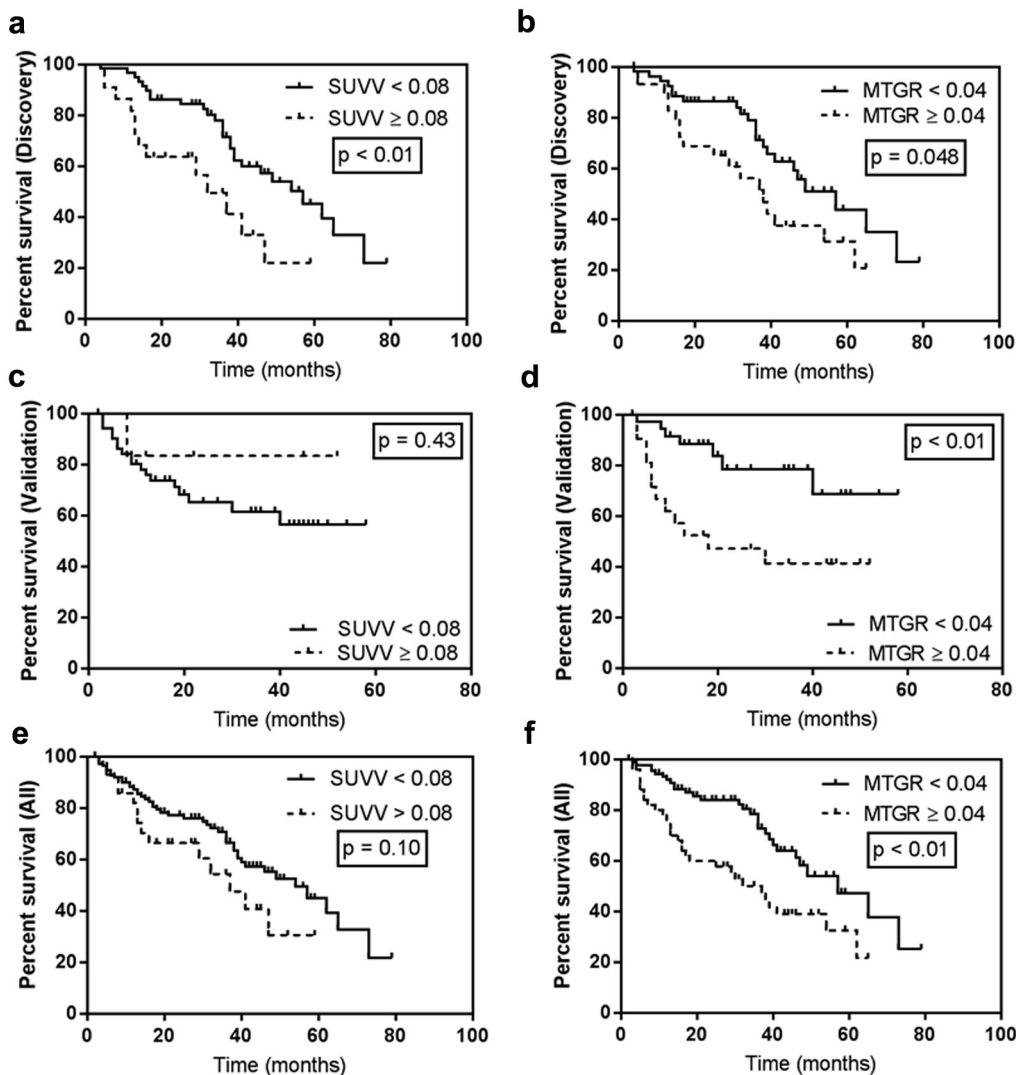

Fig. 3 Standardized uptake value velocity (SUVV; a, c, e) and metabolic tumor growth rate (MTGR; b, d, f) thresholds of 0.08 U/d and 0.04 mL/d, respectively, were identified in the discovery cohort (top row) and applied to the validation cohort (middle row) and the full cohort (bottom row). MTGR (<0.04 mL/d), but not SUVV, was prognostic of survival.

validation cohort had only 7 lesions (11.7%) with SUVV greater than 0.08 U/d. In the full cohort, patients with SUVV less than 0.08 U/d and MTGR less than 0.04 mL/d had longer median OS (54 vs 37 months, $P = .10$, and 57 vs 37 months, $P < .01$, respectively; Fig. 3).

Discussion

In patients with early stage NSCLC treated with SABR, MTGR greater than 0.04 mL/d predicted worse survival in discovery and validation cohorts. As a continuous variable, increasing MTGR trended toward prediction of distant recurrence. Although a threshold value of SUVV was not validated as prognostic of survival, increasing SUVV trended toward prognosis of LR, RR, DR, and worse survival but was not statistically significantly. These findings suggest MTGR, and possibly SUVV, may be valuable prognostic imaging biomarkers akin to doubling time and tumor grade, which are known prognostic parameters.[14-16]

Previous studies identified MTV as predictive of disease progression and death.[9-11] In apparent contradiction, single time point volumetric measurements in the present analysis, such as GTV, MTV_{Dx}, and MTV_{Sim}, were associated with decreased hazard of DR; MTV_{Sim} was also significantly associated with decreased hazard in multivariable models. However, there are at least 2 potential factors that may have contributed to these trends in our study. First, because the discovery cohort was treated chronologically earlier than the validation cohort, a treatment era effect may be a confounder. For example, 22 patients, all in the discovery cohort, were treated on a CyberKnife platform at a time during which a less accurate effective path length dose calculation algorithm was used; half of those patients (n = 11) had any disease recurrence (7 local failures). As previously reported, effective path length calculation overestimates dose to a clinically significant degree, an effect that is greater for smaller tumors.[21] Later our institution adopted tumor-volume adapted dosing of SABR in which doses are escalated to larger tumors, which mitigates the adverse prognostic impact of tumor volume[22,23]; our experience suggests similar clinical outcomes with tumor-volume adapted therapy, but there may be a control benefit with higher doses to larger tumors in this cohort. These factors and others, such as improved staging with more sensitive PET/CT scanners and increased experience with SABR over time, may have contributed to the larger number of recurrences in the discovery cohort despite smaller tumors.

We attempted to identify threshold values of MTGR and SUVV to dichotomize patients into high and low risk of death. MTGR greater than 0.04 mL/d was prognostic of worse survival. A SUVV threshold was not validated when applied to the validation set, potentially because of the small sample size; only 7 lesions in the validation cohort had SUVV greater than the threshold. Application of the thresholds to the full cohort is reported for completeness but does not represent part of the validation process. Although temporally separated cohorts can be used in validation studies,[24] this approach may be vulnerable to advances in technology used for treatment or measurement of the studied biomarker itself.

Our study had limitations. First, as a retrospective study, it was vulnerable to recall bias, convenience sampling, and unmeasured confounding. Although this study was larger than many previous retrospective studies of PET/CT-based biomarkers, statistical power was limited by the sample size. Multivariable models were limited to 4 parameters to address the small number of events. Furthermore, false discovery rate adjustment was implemented for conservative interpretation of single predictor models. Second, 9 patients with synchronous primary lung cancers (total of 21 lesions), all in the discovery cohort, were included in the study to increase power. This may increase the incidence of recurrence because of the cumulative risk from 2 independent processes or from inaccurate staging. Third, this study did not limit the PET/CT imaging device or technique; a wide variety of scanners were used, particularly for imaging at the time of diagnosis/staging. Inter- and intrascanner variability in SUV and MTV measurement may introduce variability, including the potential for negative SUVV and MTGR values. Controlling for scanner variability could have reduced SUVV and MTGR signal noise within our data and resulted in a stronger predictive effect but would have excessively limited our study cohort. On the other hand, our data represent a real-world implementation of MTGR and SUVV, and as such offer generalizability of our results to common scenarios in which staging and simulation are often performed on different scanners. Fourth, the use of multiple serial pretreatment PET/CT scans is not currently the standard of care; however, we capitalized on the routine use at our institution of PET/CT simulation for thoracic radiation therapy. The median time between diagnostic and simulation scanning in our study (56 days) is similar to that reported in other studies[25-27] and represents an area for future work-flow improvement. Our results suggest that in addition to restaging and upstaging information, simulation PET/CT may provide novel prognostic imaging biomarkers.

Conclusions

For patients with early stage NSCLC treated with SABR, an MTGR threshold of 0.04 mL/d is prognostic of survival. Our results suggest that increasing MTGR may be prognostic of distant recurrence and increasing SUVV may be prognostic of recurrence and death, but larger studies are warranted. These metrics may be promising to

study as imaging biomarkers in clinical trials of adding systemic therapy (chemotherapy or immunotherapy) to SABR to prevent regional and distant recurrence. Future studies are needed to more completely validate the prognostic and potentially predictive value of these biomarkers.

References

1. Timmerman R, Paulus R, Galvin J, et al. Stereotactic body radiation therapy for inoperable early stage lung cancer. *JAMA.* 2010;303: 1070-1076.
2. Senthi S, Lagerwaard FJ, Haasbeek CJA, Slotman BJ, Senan S. Patterns of disease recurrence after stereotactic ablative radiotherapy for early stage non-small-cell lung cancer: A retrospective analysis. *Lancet Oncol.* 2012;13:802-809.
3. Nagata Y, Hiraoka M, Shibata T, et al. Prospective trial of stereotactic body radiation therapy for both operable and inoperable T1N0M0 non-small cell lung cancer: Japan Clinical Oncology Group Study JCOG0403. *Int J Radiat Oncol Biol Phys.* 2015;93: 989-996.
4. Chang JY, Senan S, Paul MA, et al. Stereotactic ablative radiotherapy versus lobectomy for operable stage I non-small-cell lung cancer: A pooled analysis of two randomised trials. *Lancet Oncol.* 2015;16:630-637.
5. Sun B, Brooks ED, Komaki RU, et al. 7-year follow-up after stereotactic ablative radiotherapy for patients with stage I non-small cell lung cancer: Results of a phase 2 clinical trial. *Cancer.* 2017; 123:3031-3039.
6. Maziak DE, Darling GE, Inculet RI, et al. Positron emission tomography in staging early lung cancer: A randomized trial. *Ann Intern Med.* 2009;151:221-228. W-48.
7. Simpson DR, Lawson JD, Nath SK, Rose BS, Mundt AJ, Mell LK. Utilization of advanced imaging technologies for target delineation in radiation oncology. *J Am Coll Radiol JACR.* 2009;6:876-883.
8. Geiger GA, Kim MB, Xanthopoulos EP, et al. Stage migration in planning PET/CT scans in patients due to receive radiotherapy for non-small-cell lung cancer. *Clin Lung Cancer.* 2014;15:79-85.
9. Lee P, Weerasuriya DK, Lavori PW, et al. Metabolic tumor burden predicts for disease progression and death in lung cancer. *Int J Radiol Oncol Biol Phys.* 2007;69:328-333.
10. Abelson JA, Murphy JD, Trakul N, et al. Metabolic imaging metrics correlate with survival in early stage lung cancer treated with stereotactic ablative radiotherapy. *Lung Cancer Amst Neth.* 2012;78:219-224.
11. Lee P, Bazan JG, Lavori PW, et al. Metabolic tumor volume is an independent prognostic factor in patients treated definitively for non-small-cell lung cancer. *Clin Lung Cancer.* 2012;13:52-58.
12. Vesselle H, Salskov A, Turcotte E, et al. Relationship between non-small cell lung cancer FDG uptake at PET, tumor histology, and Ki-67 proliferation index. *J Thorac Oncol Off Publ Int Assoc Study Lung Cancer.* 2008;3:971-978.
13. Wang J, Mahasittiwat P, Wong KK, Quint LE, Kong F-MS. Natural growth and disease progression of non-small cell lung cancer evaluated with 18F-fluorodeoxyglucose PET/CT. *Lung Cancer Amst Neth.* 2012;78:51-56.
14. Lin C-Y, Wu Y-M, Hsieh M-H, et al. Prognostic implication of EGFR gene mutations and histological classification in patients with resected stage I lung adenocarcinoma. *PloS One.* 2017;12: e0186567.
15. Sun Z, Aubry M-C, Deschamps C, et al. Histologic grade is an independent prognostic factor for survival in non-small cell lung cancer: An analysis of 5018 hospital- and 712 population-based cases. *J Thorac Cardiovasc Surg.* 2006;131:1014-1020.
16. Usuda K, Saito Y, Sagawa M, et al. Tumor doubling time and prognostic assessment of patients with primary lung cancer. *Cancer.* 1994;74:2239-2244.
17. Yang J, Yamamoto T, Pollock S, et al. The impact of audiovisual biofeedback on 4D functional and anatomic imaging: Results of a lung cancer pilot study. *Radiother Oncol J Eur Soc Ther Radiol Oncol.* 2016;120:267-272.
18. Yu AS, von Eyben R, Yamamoto T, et al. Anatomic optimization of lung tumor stereotactic ablative radiation therapy. *Pract Radiat Oncol.* 2015;5:e607-e613.
19. Sridhar P, Mercier G, Tan J, Truong MT, Daly B, Subramaniam RM. FDG PET metabolic tumor volume segmentation and pathologic volume of primary human solid tumors. *AJR Am J Roentgenol.* 2014;202:1114-1119.
20. Camp RL, Dolled-Filhart M, Rimm DL. X-tile: A new bioinformatics tool for biomarker assessment and outcome-based cut-point optimization. *Clin Cancer Res Off J Am Assoc Cancer Res.* 2004;10:7252-7259.
21. Liu MB, Eclov NCW, Trakul N, et al. Clinical impact of dose overestimation by effective path length calculation in stereotactic ablative radiation therapy of lung tumors. *Pract Radiat Oncol.* 2013; 3:294-300.
22. Brown JM, Diehn M, Loo BW. Stereotactic ablative radiotherapy should be combined with a hypoxic cell radiosensitizer. *Int J Radiat Oncol Biol Phys.* 2010;78:323-327.
23. Trakul N, Chang CN, Harris J, et al. Tumor volume-adapted dosing in stereotactic ablative radiotherapy of lung tumors. *Int J Radiat Oncol Biol Phys.* 2012;84:231-237.
24. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): The TRIPOD statement. *Ann Intern Med.* 2015;162:55-63.
25. Murai T, Shibamoto Y, Baba F, et al. Progression of non-small-cell lung cancer during the interval before stereotactic body radiotherapy. *Int J Radiat Oncol Biol Phys.* 2012;82:463-467.
26. O'Rourke N, Edwards R. Lung cancer treatment waiting times and tumour growth. *Clin Oncol R Coll Radiol G B.* 2000;12:141-144.
27. Mohammed N, Kestin LL, Grills IS, et al. Rapid disease progression with delay in treatment of non−small-cell lung cancer. *Int J Radiat Oncol.* 2011;79:466-472.