Predictive value of the heterogeneity index in assessing treatment response using ^{18}F-FDG PET/CT in locally advanced and advanced lung cancers

Yiğithan Okar[1] MD,
Burcu Esen Akkaş[2] MD,
Reyhan Toyran[3] MD,
Muammer Urhan[3] MD

1. Umraniye Training and Research Hospital, Istanbul, Turkiye.
2. Basaksehir Cam and Sakura City Hospital, Istanbul, Turkiye.
3. Sultan 2. Abdülhamid Han Training and Research Hospital, Istanbul, Turkiye.

Keywords: ^{18}F-FDG PET/CT
- Heterogeneity index
- Metabolic tumor volume
- Lung cancer
- Treatment response

Corresponding author:
Yiğithan Okar MD,
Umraniye Training and Research Hospital, Kısıklı Mahallesi, Alem dağ Caddesi, No: 31, Üsküdar, Istanbul, Turkiye
Tel: 0536 996 69 52
yigithanokar@gmail.com

Received:
11 April 2026
Accepted revised:
3 August 2026

Abstract

Objective: Intratumoral metabolic heterogeneity may influence treatment response in lung cancer, and the heterogeneity index (HI), derived from threshold-based metabolic tumor volume (MTV) measure-ments, has been proposed as a simple fluorine-18-fluorodeoxyglucose (^{18}F-FDG) positron emission tomography/computed tomography (PET/CT)-based parameter to estimate this variability. This retrospective single-center study evaluated the predictive value of HI for treatment response in patients with locally advanced and advanced lung cancer undergoing chemoradiotherapy and compared its performance with conventional PET-derived volumetric parameters. **Subjects and Methods:** Forty-four patients with stage III–IV lung cancer who underwent pretreatment ^{18}F-FDG PET/CT followed by chemoradiotherapy were retrospectively analyzed. Positron emission tomography-derived parameters, including maximum standardized uptake value (SUVmax), mean standardized uptake value (SUVmean), MTV, total lesion glycolysis (TLG), and HI, were calculated from baseline scans. Heterogeneity index was derived from the slope of linear regression using MTV values measured at multiple threshold levels. Treatment response was classified as regression or progression based on follow-up imaging according to positron emission tomography response criteria in solid tumors (PERCIST). Receiver operating characteristic (ROC) analysis was used to evaluate the predictive performance of MTV and HI, and multivariate logistic regression was performed to identify independent predictors of disease progression. **Results:** Of the 44 patients, 19 (43.2%) showed regression and 25 (56.8%) showed progression. Metabolic tumor volume and HI were significantly higher in patients with progression (P=0.047 and P=0.040, respectively). Receiver operating characteristic analysis demonstrated modest discriminative performance for MTV (area under the curve [AUC]: 0.677; 95% confidence interval [CI]: 0.519–0.810) and HI (AUC: 0.682; 95% CI: 0.525–0.814), with high sensitivity but low specificity for both parameters. Heterogeneity index showed an extremely strong correlation with MTV (r= 0.997), indicating substantial overlap between the two measures. In multivariate analysis, only advanced tumor stage (T3-T4) remained an independent predictor of progression (odds ratio [OR]: 6.801; 95% CI: 1.271–36.398; P=0.025), whereas HI did not retain independent significance. **Conclusions:** These findings indicate that HI has limited but statistically significant predictive value for treatment response in locally advanced and advanced lung cancer and performs similarly to MTV. Given its extreme collinearity with MTV, HI should be interpreted primarily as a derivative parameter reflecting tumor burden rather than as an independent heterogeneity biomarker. Nevertheless, because of its simple calculation and broad availability, HI may serve as a practical adjunct metabolic parameter in clinical assessment, although larger prospective studies with external validation are needed to clarify its incremental clinical utility.

Hell J Nucl Med 2026; 29(2): 94-103 *Epub ahead of print: 6 August 2026* *Published online: 31 August 2026*

Introduction

Lung cancer is the third most common malignancy after breast cancer in women and prostate cancer in men, and it remains the leading cause of cancer-related mortality worldwide [1]. Owing to its heterogeneous etiopathogenesis, the disease exhibits variable clinical behavior and treatment response; genetic factors are thought to play a significant role in this variability [2]. The use of molecular biomarker-based targeted therapies has led to marked improvements in survival outcomes.

Accurate staging is critical for determining treatment and prognosis in lung cancer, with approximately 15% of cases classified as small-cell lung cancer (SCLC) and 85% as non-small-cell lung cancer (NSCLC) [3]. In TNM staging, fluorine-18-fluorodeoxyglucose (^{18}F-FDG) positron emission tomography/computed tomography (PET/CT) has demonstrated superiority over conventional methods in the evaluation of solitary pulmonary

nodules (SPN), nodal staging, radiotherapy planning, and detection of recurrence [4, 5].

The high rate of post-chemotherapy progression in both NSCLC and SCLC underscores the importance of accurate assessment of treatment response [6]. Computed tomography-based response evaluation criteria in solid tumors (RECIST) have notable limitations in lesions with irregular borders, with higher rates of misclassification, highlighting the value of PET-based evaluation [7, 8]. Since metabolic changes precede anatomical alterations, ^{18}F-FDG PET/CT is superior for the early detection of progression [9]. The superiority of positron emission tomography response criteria in solid tumors (PERCIST) over RECIST has been demonstrated in several studies [10, 11]. Additionally, PET/CT has proven superior to conventional techniques in detecting recurrence [12, 13].

Although maximum standardized uptake value (SUVmax), metabolic tumor volume (MTV), and total lesion glycolysis (TLG) are fundamental parameters for evaluating treatment response, SUVmax may not adequately reflect intratumoral metabolic heterogeneity [14, 15]. Metabolic tumor volume and TLG have been reported to provide a more accurate representation of tumor burden and metabolic activity [16].

Intratumoral heterogeneity, reflecting cellular, genetic, and metabolic variations, has been associated with treatment response and prognosis in several malignancies [17]. The ^{18}F-FDG PET/CT-based linear regression slope method provides a practical and quantitative approach for estimating heterogeneity [18]. Previous studies in gastric and colorectal cancers have suggested that this index may have prognostic value [18-20]. Similarly, studies across different tumor types have reported potentially useful associations between the heterogeneity index and clinical outcomes, although its incremental value relative to conventional PET parameters has not been consistent across settings [21-24].

This study aimed to evaluate the relationship between baseline PET/CT-derived tumor heterogeneity and treatment response in lung cancer patients and to compare the performance of the heterogeneity index with conventional PET parameters including SUVmax, MTV, and TLG.

Subjects and Methods

Ethics approval

Ethical approval for the study was obtained from the Ethics Committee of T.C. SBÜ Hamidiye Scientific Research Board on 14 March 2024, with decision number 3/8.

Case selection

A total of 184 newly diagnosed lung cancer patients who underwent ^{18}F-FDG PET/CT between 2018 and 2023 were retrospectively evaluated. Pathology, radiology, and oncology records were accessed through the hospital information system. Fifty-one patients with unavailable pathology data, 67 patients without follow-up PET imaging, and 22 patients whose follow-up scans were acquired at external centers were excluded. The study included inoperable Stage III–IV SCLC and NSCLC patients who received chemoradiotherapy. Baseline PET/CT scans were available for 44 patients, and all follow-up imaging studies were complete. According to PERCIST criteria, 19 patients (43.2%) were classified as regressive and 25 patients (56.8%) as progressive. Patients received standard treatment in accordance with National Comprehensive Cancer Network (NCCN) guidelines, and PET/CT follow-up imaging was performed at 3-6-month intervals. The patient selection flowchart is presented in Figure 1.

Method

Maximum SUV, mean standardized uptake value (SUVmean), MTV, TLG, and the heterogeneity index were calculated from the baseline staging PET/CT scans. Heterogeneity index was calculated exclusively from the primary lung tumor to ensure the evaluation of a uniform and anatomically consistent target across all patients. Metastatic lesions were not included because their number, size, anatomical distribution, and ^{18}F-FDG avidity varied considerably among patients, which could have introduced substantial lesion-selection variability and limited standardized interpatient comparisons. After identifying primary tumor boundaries, MTV values corresponding to 40%, 60%, and 80% threshold levels were measured. For each primary tumor, the volume of interest (VOI) was delineated manually by two nuclear medicine physicians with at least 5 years of experience, blinded to clinical data and outcomes. Volume of interest boundaries were drawn with particular attention to avoid inclusion of surrounding physiological ^{18}F-FDG uptake (heart, major vessels, mediastinum, inflammatory tissues, etc.). A graph representing the MTV-threshold relationship was generated in Microsoft Excel. Linear regression slope values were calculated, and the heterogeneity index was defined as the negative value of this slope. Because HI was calculated directly from MTV measurements obtained at multiple relative thresholds, it is mathematically dependent on MTV; therefore, a strong association and potential collinearity between these two parameters are inherently expected.

Patients were evaluated after 6-8 hours of fasting, with blood glucose levels required to be below 180mg/dL. All patients received 0.15mCi/kg of ^{18}F-FDG intravenously, and imaging was performed 60 minutes later from the skull base to mid-thigh. Attenuation correction was performed using simultaneous low-dose CT. All scans were acquired on the same PET/CT system (GE Discovery 600) in 3D mode using 7-9 bed positions. Images were reconstructed using ordered-subset expectation maximization (OSEM). Measurements were performed on the GE Volume Viewer workstation, and MTV and TLG values were obtained using threshold levels of 40%, 60%, and 80%.

Statistical analysis

Categorical variables were presented as frequencies and percentages and analyzed using Pearson's chi-square or Fisher's exact test. The distribution of continuous variables was assessed with the Shapiro-Wilk test. Normally distributed variables were reported as mean±standard deviation (SD), whereas non-normally distributed variables were expressed as median with interquartile range (IQR). Comparisons bet-

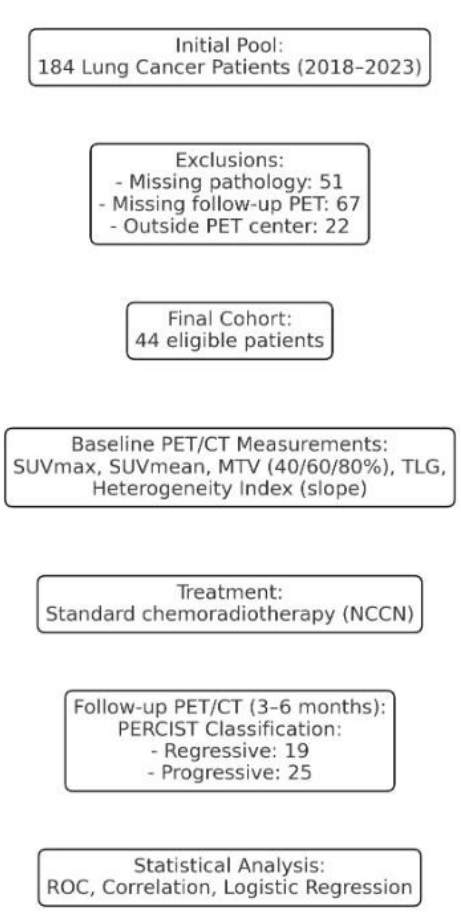

Figure 1. Patient selection flowchart.

ween groups were performed using the independent t-test or the Mann-Whitney U test. The predictive performance of MTV and the heterogeneity index for progression was evaluated using receiver operating characteristic (ROC) analysis, and area under the curve (AUC), cut-off values, sensitivity, specificity, and 95% CI were reported. Optimal cut-off values were determined using the Youden index. Area under the curve comparisons were performed with the DeLong method [25]. Correlation analysis was conducted using Spearman's test, and independent risk factors were assessed through multivariate logistic regression. All analyses were performed using IBM SPSS 23.0, and $P<0.05$ was considered statistically significant.

Results

The demographic and clinical characteristics of the 44 patients included in the study are summarized in Table 1. The mean age was 63.18±8.31 years, and 72.7% of the patients were male. Pathologically, 72.7% of cases were diagnosed with NSCLC, while 27.3% had SCLC. Regarding T stage, the majority of patients were classified as T3 (27.3%) or T4 (43.2%). In terms of nodal involvement, 45.5% had N2 and 25.0% had N3 disease. Overall, 72.7% of patients were M0 and 27.3% were M1, corresponding to 72.7% Stage III and 27.3% Stage IV disease.

Positron emission tomography-derived parameters were non-normally distributed and therefore reported as median with IQR. The median values were as follows: SUVmax 11.5 (IQR: 8.15-15.25), SUVmean 6.7 (IQR: 4.8-8.55), MTV 43.52 (IQR: 24.21-90.47), TLG 272.56 (IQR: 90.26-704.33), and heterogeneity index 1.00 (IQR: 0.53-2.13). At the end of the evaluation, 19 patients (43.2%) demonstrated regression, while 25 patients (56.8%) showed progression.

Comparisons based on progression status revealed no significant differences between groups with respect to age, sex, pathology, N stage, metastasis, or overall stage ($P>0.05$). Patients exhibiting regression had higher proportions of T1-T2 disease, whereas those in the progressive group showed a markedly higher proportion of T3-T4 tumors ($P=0.037$). No significant differences were observed in SUVmax or SUVmean ($P=0.740$ and $P=0.981$, respectively). Metabolic tumor volume values were significantly higher in the progressive group (55.84; IQR: 31.2-115) compared with the regression group (36.28; IQR: 9.49-79.51) ($P=0.047$). Similarly, the heterogeneity index was higher in the progressive group (1.28 [IQR: 0.68-2.78]) than in the regressive group (0.85 [IQR: 0.20-1.92]; $P=0.040$). Total lesion glycolysis values also tended to be higher in progressive cases (313.749 [IQR: 135.466-725.514] vs. 226.549 [IQR: 72.825-683.136]), although this difference did

Table 1. *Baseline characteristics of the study population.*

Variable	Total (n=44)
Age (years), mean±SD	63.18±8.31
Sex, n (%)	
Female	12 (27.3)
Male	32 (72.7)
Histology, n (%)	
NSCLC	32 (72.7)
SCLC	12 (27.3)
T stage, n (%)	
T1	5 (11.4)
T2	8 (18.2)
T3	12 (27.3)
T4	19 (43.2)
N stage, n (%)	
N0	11 (25.0)
N1	2 (4.5)
N2	20 (45.5)
N3	11 (25.0)
M stage, n (%)	
M0	32 (72.7)
M1	12 (27.3)
PET metrics, median (IQR)	
SUVmax	11.5 (8.15-15.25)
SUVmean	6.7 (4.8-8.55)
MTV	43.52 (24.21-90.47)
TLG	272.56 (90.26-704.33)
Heterogeneity index (HI)	1.00 (0.53-2.13)

NSCLC, non-small-cell lung cancer; SCLC, small-cell lung cancer; SUVmax, maximum standardized uptake value; SUVmean, mean standardized uptake value; MTV, metabolic tumor volume; TLG, total lesion glycolysis; HI, heterogeneity index; IQR, interquartile range.

Table 2. *Comparison between regressive and progressive groups.*

Variable	Regressive (n=19)	Progressive (n=25)	P-value
Age (years), mean±SD	62.6±9.7	63.6±7.2	0.680
Male sex, n (%)	14 (73.7)	18 (72.0)	0.901
SCLC histology, n (%)	4 (21.1)	8 (32.0)	0.419
T3–T4 stage, n (%)	9 (47.4)	22 (88.0)	0.037
N2–N3 stage, n (%)	15 (78.9)	16 (64.0)	0.734
M1 disease, n (%)	4 (21.1)	8 (32.0)	0.419
PET metrics, median (IQR)			
SUVmax	12.0 (6.6-16.1)	11.5 (9.6-14.0)	0.740
SUVmean	7.6 (3.9-9.8)	6.7 (5.5-7.8)	0.981
MTV	36.28 (9.49-79.51)	55.84 (31.2-115.0)	0.047
TLG	226.549 (72.825-683.136)	313.749 (135.466-725.514)	0.139
Heterogeneity index (HI)	0.85 (0.20-1.92)	1.28 (0.68-2.78)	0.040

SCLC, small-cell lung cancer; MTV, metabolic tumor volume; TLG, total lesion glycolysis; HI, heterogeneity index; IQR, interquartile range.

not reach statistical significance (P=0.139).

In the ROC analysis for MTV and the heterogeneity index, AUC values were calculated as 0.677 (95% CI: 0.519-0.810; P= 0.041) and 0.682 (95% CI: 0.525-0.814; P=0.034), respectively (Table 3 and Figure 2). The optimal cut-off value for MTV was >22.3 (sensitivity 92%; specificity 47.4%), while the optimal cut-off for the heterogeneity index was >0.45 (sensitivity 96%; specificity 47.4%). The discriminative performances of the two parameters were similar, with no significant difference between their AUC values (P>0.05).

The results of the Spearman correlation analysis between the heterogeneity index and other study parameters are presented in Table 4. The heterogeneity index demonstrated a very strong positive correlation with MTV (r=0.997; P< 0.001) and TLG (r=0.897; P<0.001) (Figures 3 and 4). A weak but statistically significant positive correlation was observed with SUVmax (r=0.339; P=0.024) (Figure 5). The correlation with SUVmean was weak and did not reach statistical signifi-

Table 3. *ROC analysis for predicting progression.*

Parameter	AUC (95% CI)	P-value	Cut-off value	Sensitivity (%)	Specificity (%)
MTV	0.677 (0.519-0.810)	0.041	>22.3	92.0	47.4
Heterogeneity index (HI)	0.682 (0.525-0.814)	0.034	>0.45	96.0	47.4

ROC, receiver operating characteristic; AUC, area under the curve; CI, confidence interval; MTV, metabolic tumor volume; HI, heterogeneity index.

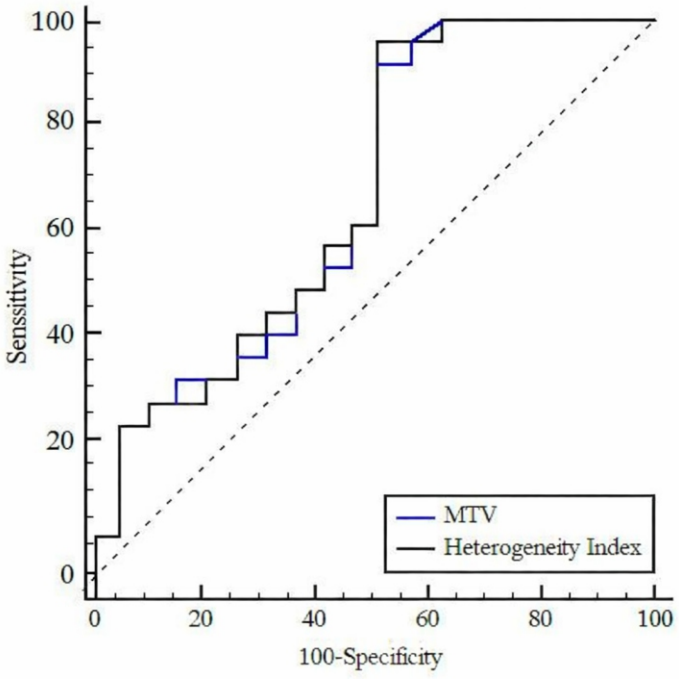

Figure 2. Receiver operating characteristic curves for MTV and the heterogeneity index in distinguishing progression among patients.

cance (P=0.052). No significant correlation was found between age and the heterogeneity index (P=0.482).

Given the extremely strong correlation observed between MTV and the heterogeneity index, multicollinearity diagnostics were performed. Variance inflation factor (VIF) analysis demonstrated very high collinearity between these two parameters (VIF ≈293). Therefore, MTV was not included in the multivariate logistic regression model in order to avoid multicollinearity. In multivariate logistic regression analysis including T stage and the heterogeneity index, only advanced tumor stage (T3-T4) remained independently associated with disease progression (OR: 6.801; 95% CI: 1.271-36.398; P=0.025) (Table 5). The heterogeneity index did not demonstrate independent predictive value (P=0.647). The multivariate model included two predictors, resulting in an events-per-variable (EPV) ratio of 12.5, which is above the commonly recommended threshold of 10.

Discussion

The concept of tumor heterogeneity refers to the genetic, biological, and molecular alterations that arise during tumor growth, leading to differences in proliferation rate, invasive potential, drug sensitivity, and prognosis among newly formed daughter cells after multiple cycles of cellular division. This phenomenon can be regarded as one of the hallmark characteristics of malignant tumors [17]. In ^{18}F-FDG PET/CT studies, intratumoral heterogeneity is commonly assessed using classical semi-quantitative parameters such as SUV, MTV, and TLG, as well as through three-dimensional texture analyses performed with dedicated software programs such as LIFEx.

Tumor heterogeneity has emerged as a major challenge in personalized oncology. In nuclear medicine practice, this challenge is reflected in treatment resistance or progressive disease. Accordingly, characterization of tumor heterogeneity and its potential predictive value has recently gained considerable attention in oncologic PET/CT research.

Pyka et al. (2015) [26] demonstrated that, among 45 medically inoperable patients with early-stage NSCLC, tumors

Table 4. *Correlation of heterogeneity index with other parameters.*

Variable	r	P-value
Age	-0.109	0.482
SUVmax	0.339	0.024
SUVmean	0.295	0.052
MTV	0.997	<0.001
TLG	0.897	<0.001

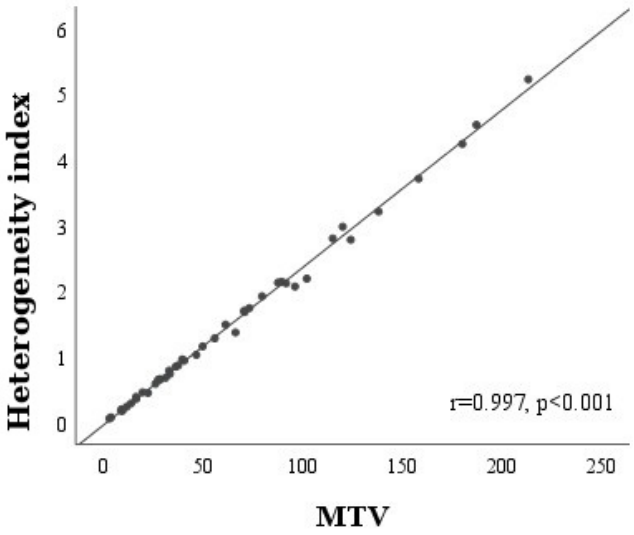

Figure 3. Correlation between the heterogeneity index and MTV.

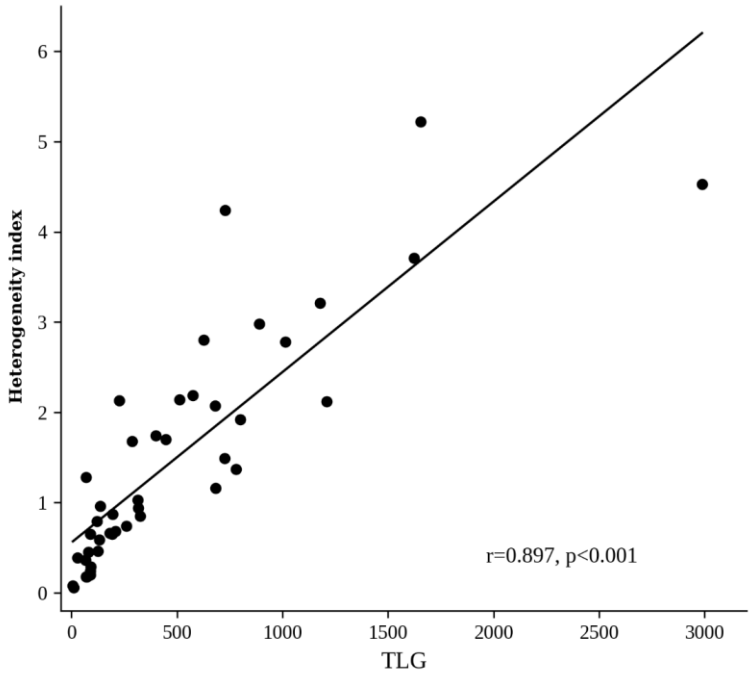

Figure 4. Correlation between the heterogeneity index and TLG.

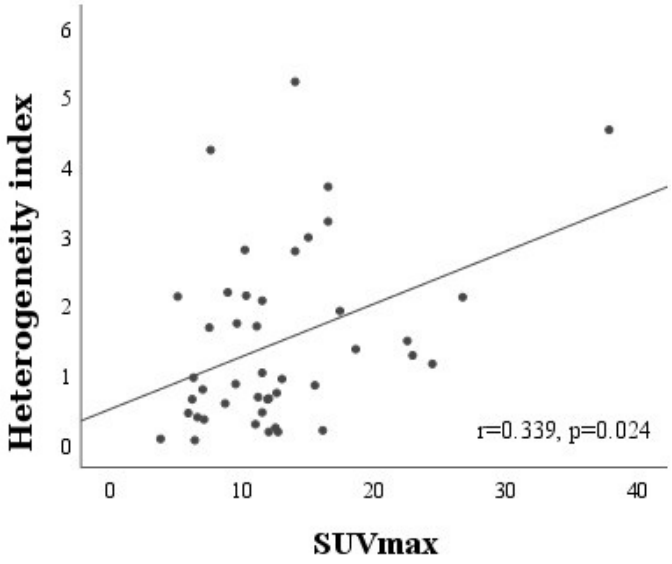

Figure 5. Correlation between the heterogeneity index and SUVmax.

Table 5. *Multivariate logistic regression.*

Variable	OR (95% CI)	P-value
T stage (T3-T4 vs T1-T2)	6.801 (1.271-36.398)	0.025
Heterogeneity index (HI)	1.155 (0.623-2.140)	0.647

OR, odds ratio; CI, confidence interval; HI, heterogeneity index.

with negative ^{18}F-FDG uptake patterns and those with high SUVmax values exhibited an increased risk of local recurrence on pretreatment ^{18}F-FDG PET/CT imaging. Their structural tumor analyses further showed that PET texture analysis was an independent factor for disease-specific survival. Similarly, Cook et al. (2015) [27] investigated tumor heterogeneity and classical semi-quantitative PET parameters in 47 NSCLC patients undergoing erlotinib therapy and demonstrated a positive correlation between lower levels of heterogeneity and treatment response. Although various software tools such as LIFEx facilitate advanced tumor texture analysis, their routine clinical use remains limited because of restricted accessibility, lack of standardized heterogeneity criteria, and technical complexity.

Consequently, simpler methods based on conventional PET parameters have been developed to estimate tumor heterogeneity. The linear regression slope method used in our study is one such approach; it is inexpensive, widely accessible, easily applicable on most workstations, and allows a simplified approximation of intratumoral metabolic variability based on volumetric PET measurements. Unlike HI, texture-based metrics characterize voxel-level patterns and spatial relationships within the tumor and may therefore provide a more detailed description of intratumoral ^{18}F-FDG uptake heterogeneity [26, 27, 30]. However, texture features may be influenced by image acquisition and reconstruction parameters, tumor segmentation, intensity discretization, tumor volume, and the software used for analysis, and they generally require dedicated radiomics platforms and methodological standardization. In contrast, HI can be calculated from routinely available threshold-based MTV measurements and is therefore simpler and more accessible. Nevertheless, HI represents a relatively coarse and volume-dependent approximation of heterogeneity and should not be considered equivalent to more comprehensive texture-based analysis.

To our knowledge, Kidd et al. (2008) [28] were the first to apply the linear regression slope method to quantify tumor heterogeneity in cervical cancer. Their study demonstrated that intratumoral ^{18}F-FDG heterogeneity calculated before treatment had predictive value for lymph node metastasis risk, response to radiotherapy, and pelvic recurrence, and that the heterogeneity index was superior to SUVmax.

Several mechanisms have been proposed to explain the association between metabolic heterogeneity and prognosis. Fluorine-18-FDG uptake reflects factors such as GLUT and hexokinase expression, vascularization, and tumor cell proliferation [29]. Increased metabolic variability may correspond to pronounced neovascularization, a characteristic of more aggressive tumors. Tumor heterogeneity may also be attributable to hypoxia, which is a known negative prognostic factor [30]. Vansteenkiste et al. (1999) [31] reported

that a primary tumor SUVmax ≥7 was associated with poor prognosis in 125 patients with NSCLC. Similarly, Tsutani et al. (2011) [32] demonstrated the strong prognostic significance of SUVmax in lung adenocarcinoma. However, because SUVmax represents only a focal point of metabolic activity rather than the overall metabolic behavior of the tumor, it may be insufficient for evaluating intratumoral metabolic variability.

Huang et al. (2012) [33] showed that the heterogeneity index was superior to classical semi-quantitative PET parameters such as SUVmax and MTV for predicting disease progression in newly diagnosed nasopharyngeal carcinoma. Tixier et al. (2011) [34] demonstrated in 41 esophageal cancer patients that pretreatment heterogeneity, assessed using texture features, was superior to SUVmax for predicting response to chemoradiotherapy. Kim et al. (2017) [35] found that in 93 patients with pancreatic ductal adenocarcinoma who underwent preoperative ^{18}F-FDG PET/CT, the heterogeneity index showed improved predictive performance compared with SUVmax and other semi-quantitative PET parameters. Consistent with previous studies, SUVmax was not significantly associated with treatment response in our cohort of stage III–IV lung cancer patients receiving chemoradiotherapy. Volume-based PET parameters such as MTV and TLG have been shown to outperform SUVmax in prognostic assessment in lung cancer. For instance, Zhang et al. (2013) [36] demonstrated in a cohort of 104 stage I and IV NSCLC patients that pretreatment MTV and TLG were superior to SUVmax for predicting prognosis.

In the present study, the heterogeneity index demonstrated modest discriminative performance, with results largely comparable to MTV rather than clearly superior to established volumetric PET parameters. The ROC analyses revealed similar AUC values for HI and MTV, suggesting that both parameters likely capture overlapping aspects of tumor burden and metabolic activity. Although ROC analysis demonstrated statistically significant AUC values for both MTV and HI, their clinical utility should be interpreted with caution. The specificity of both parameters was only 47.4%, corresponding to a substantial false-positive rate among patients without progression. Therefore, use of either parameter in isolation could incorrectly classify a considerable proportion of regressive patients as being at high risk of progression and potentially lead to unnecessary intensified surveillance or treatment-related evaluation. Their high sensitivity suggests that MTV and HI may contribute to broader risk-stratification models; however, their low specificity precludes their use as standalone clinical decision-making tools.

The extremely strong correlation between HI and MTV ® 0.997) should be interpreted in light of the mathematical construction of HI. Because HI is directly derived from changes in MTV across multiple relative thresholds, a strong association and substantial collinearity between the two parameters are inherently expected rather than representing an unexpected biological finding. Accordingly, HI may largely behave as a derived volumetric parameter reflecting tumor burden and threshold-dependent volume variation rather than as a fully independent biological marker of intratumoral heterogeneity. Therefore, HI should be interpreted as a complementary metric derived from MTV measurements rather than an independent heterogeneity biomarker.

Although the heterogeneity index was significantly associated with progression in univariate analysis, it did not remain an independent predictor in multivariate analysis including T stage and the heterogeneity index. Nevertheless, despite its lack of independent predictive value, HI may still have practical adjunctive value because it can be readily calculated from routinely available MTV measurements and may provide supplementary risk-stratification information when interpreted alongside tumor stage and other clinical parameters, rather than as a standalone biomarker. Multivariate logistic regression demonstrated that patients with T3-T4 tumors had a significantly higher risk of disease progression. Advanced T stages may inherently reflect greater tumor burden and biological aggressiveness, which may also explain their stronger association with progression in our cohort. Furthermore, the predominance of T3-T4 tumors in this cohort, although expected given the inclusion of patients with locally advanced and advanced disease, may have limited the ability of the multivariable model to demonstrate incremental predictive information from HI beyond T stage.

Most prior studies have predominantly focused on NSCLC subtypes. In our study, histopathological subtype was not a significant determinant of treatment response. Another interesting observation was that the presence of distant metastasis at diagnosis did not significantly influence treatment response. In the present study, HI was calculated exclusively from the primary lung tumor to provide a standardized and anatomically consistent target across the entire cohort. Inclusion of metastatic lesions would have introduced additional variability because the number, size, location, and metabolic activity of metastatic deposits differed substantially among patients. Nevertheless, metabolic characteristics may vary between the primary tumor and metastatic sites [38]. Therefore, primary-tumor HI may not fully represent interlesional or whole-body tumor heterogeneity, particularly in patients with metastatic disease.

Our study has several limitations, including the relatively small and heterogeneous sample, the limited number of progression events, the retrospective design, and manual VOI delineation. Furthermore, because HI was derived only from the primary tumor, the analysis did not account for interlesional metabolic heterogeneity between primary and metastatic sites. In addition, texture-based heterogeneity features were not evaluated, precluding a direct within-cohort comparison between HI and more comprehensive radiomic metrics. The low specificity of both HI and MTV resulted in a substantial false-positive rate, limiting their suitability as standalone predictors of progression. The optimal cut-off values were derived and evaluated in the same cohort, which may have resulted in optimistic estimates of their predictive performance. No independent external cohort was available to assess the reproducibility, generalizability, or clinical applicability of these thresholds. Moreover, because HI is mathematically derived from threshold-based MTV measurements and showed extreme collinearity with MTV in our dataset, its interpretation as an independent heterogeneity biomarker remains limited. Therefore, the reported cut-off values should be considered exploratory rather than ready for clinical implementation. Larger prospective multi-

center studies using standardized acquisition, reconstruction, and segmentation protocols are needed to externally validate these findings and determine whether HI provides clinically meaningful information beyond established volumetric PET parameters.

In conclusion, in patients with inoperable stage III-IV lung cancer, the heterogeneity index showed modest predictive performance comparable to that of MTV. Given its extreme collinearity with MTV and lack of independent significance in multivariable analysis, HI should be interpreted cautiously as a derived parameter reflecting tumor burden rather than as an independent heterogeneity biomarker. Nevertheless, because of its simple calculation from routinely available MTV measurements and broad accessibility, HI may still serve as a practical adjunct for risk stratification when interpreted alongside tumor stage and other clinical parameters, although it should not be used as a standalone predictor. Further prospective studies are needed to determine whether HI provides clinically meaningful value beyond established volumetric PET parameters.

The authors declare that they have no conflicts of interest.

Bibliography

1. Heron M. Deaths: leading causes for 2019. *Natl Vital Stat Rep* 2021; 70: 1-114.
2. Arnold BN, Thomas DC, Rosen JE et al. Lung cancer in the very young: treatment and survival in the National Cancer Data Base. *J Thorac Oncol* 2016; 11: 1121-31.
3. Semenova EA, Nagel R, Berns A. Origins, genetic landscape, and emerging therapies of small cell lung cancer. *Genes Dev* 2015; 29: 1447-62.
4. Greco C, Rosenzweig K, Cascini GL et al. Current status of PET/CT for tumour volume definition in radiotherapy treatment planning for non-small cell lung cancer (NSCLC). *Lung Cancer* 2007; 57: 125-34.
5. Devaraj A, Cook GJR, Hansell DM. PET/CT in non-small cell lung cancer staging-promises and problems. *Clin Radiol* 2007; 62: 97-108.
6. Nishino M, Hatabu H, Johnson BE, McLoud TC. State of the art: response assessment in lung cancer in the era of genomic medicine. *Radiology* 2014; 271: 6-27.
7. Erasmus JJ, Gladish GW, Broemeling L et al. Interobserver and intraobserver variability in measurement of non-small-cell carcinoma lung lesions: implications for assessment of tumor response. *J Clin Oncol* 2003; 21: 2574-82.
8. William WN Jr, Pataer A, Kalhor N et al. Computed tomography RECIST assessment of histopathologic response and prediction of survival in patients with resectable non-small-cell lung cancer after neoadjuvant chemotherapy. *J Thorac Oncol* 2013; 8: 222-8.
9. Dunlap NE, Yang W, McIntosh A et al. Computed tomography-based anatomic assessment overestimates local tumor recurrence in patients with mass-like consolidation after stereotactic body radiotherapy for early-stage non-small cell lung cancer. *Int J Radiat Oncol Biol Phys* 2012; 84: 1071-7.
10. Mac Manus MP, Hicks RJ, Matthews JP et al. Positron emission tomography is superior to computed tomography scanning for response-assessment after radical radiotherapy or chemoradiotherapy in patients with non-small-cell lung cancer. *J Clin Oncol* 2003; 21: 1285-92.
11. Kremer R, Peysakhovich Y, Dan LF et al. ^{18}F-FDG PET/CT for assessing the resectability of NSCLC patients with N2 disease after neoadjuvant therapy. *Ann Nucl Med* 2016; 30: 114-21.
12. He YQ, Gong HL, Deng YF, Li WM. Diagnostic efficacy of PET and PET/CT for recurrent lung cancer: a meta-analysis. *Acta Radiol* 2014; 55: 309-17.
13. Choi SH, Kim YT, Kim SK et al. Positron emission tomography-computed tomography for postoperative surveillance in non-small cell lung cancer. *Ann Thorac Surg* 2011; 92: 1826-32.
14. Na F, Wang J, Li C et al. Primary tumor standardized uptake value measured on ^{18}F-fluorodeoxyglucose positron emission tomography is of prediction value for survival and local control in non-small-cell lung cancer receiving radiotherapy: meta-analysis. *J Thorac Oncol* 2014; 9: 834-42.
15. Downey RJ, Akhurst T, Gonen M et al. Preoperative ^{18}F-fluorodeoxyglucose-positron emission tomography maximal standardized uptake value predicts survival after lung cancer resection. *J Clin Oncol* 2004; 22: 3255-60.
16. Cheng G, Huang H. Prognostic value of ^{18}F-fluorodeoxyglucose PET/CT in non-small-cell lung cancer. *PET Clin* 2018; 13: 59-72.
17. Marusyk A, Polyak K. Tumor heterogeneity: causes and consequences. *Biochim Biophys Acta* 2010; 1805: 105-17.
18. Wang J, Yu X, et al. Predictive value of ^{18}F-FDG PET/CT multi-metabolic parameters and tumor metabolic heterogeneity in the prognosis of gastric cancer. *J Cancer Res Clin Oncol* 2023; 149: 14535-47.
19. Liu X, Xiang K et al. Prognostic value of intratumor metabolic heterogeneity parameters on ^{18}F-FDG PET/CT for patients with colorectal cancer. *Contrast Media Mol Imaging* 2022; 2022: 2586245.
20. Liu G et al. Intra-tumor metabolic heterogeneity of gastric cancer on ^{18}F-FDG PET/CT indicates patient survival outcomes. *Clin Exp Med* 2021; 21: 129-38.
21. Kang SR, Song HC, Byun BH et al. Intratumoral metabolic heterogeneity for prediction of disease progression after concurrent chemoradiotherapy in patients with inoperable stage III non-small-cell lung cancer. *Nucl Med Mol Imaging* 2014; 48: 16-25.
22. Lee M, Lee H, Cheon GJ et al. Prognostic value of preoperative intratumoral FDG uptake heterogeneity in patients with epithelial ovarian cancer. *Eur Radiol* 2017; 27: 16-23.
23. Kimura M, Kato I, Ishibashi K et al. The prognostic significance of intratumoral heterogeneity of ^{18}F-FDG uptake in patients with oral cavity squamous cell carcinoma. *Eur J Radiol* 2019; 114: 99-104.
24. Chung HH, Kang SY, Ha S et al. Prognostic value of preoperative intratumoral FDG uptake heterogeneity in early stage uterine cervical cancer. *J Gynecol Oncol* 2016; 27: e15.
25. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988; 44: 837-45.
26. Pyka T, Bundschuh RA et al. Textural features in pre-treatment ^{18}F-FDG-PET/CT are correlated with risk of local recurrence and disease-specific survival in early stage NSCLC patients receiving primary stereotactic radiation therapy. *Radiat Oncol* 2015; 10: 100.
27. Cook GJR, O'Brien ME, Siddique M et al. Non-small cell lung cancer treated with erlotinib: heterogeneity of ^{18}F-FDG uptake at PET- association with treatment response and prognosis. *Radiology* 2015; 276: 883-93.
28. Kidd EA, Grigsby PW. Intratumoral metabolic heterogeneity of cervical cancer. *Clin Cancer Res* 2008; 14: 5236-41.
29. Rajendran JG, Schwartz DL, O'Sullivan J et al. Tumor hypoxia imaging with ^{18}F-fluoromisonidazole positron emission tomography in head and neck cancer. *Clin Cancer Res* 2006; 12: 5435-41.
30. Bundschuh RA, Dinges J, Neumann L et al. Textural parameters of tumor heterogeneity in ^{18}F-FDG PET/CT for therapy response assessment and prognosis in patients with locally advanced rectal cancer. *J Nucl Med* 2014; 55: 891-7.
31. Vansteenkiste JF, Stroobants SG, Dupont PJ et al. Prognostic importance of the standardized uptake value on ^{18}F-fluoro-2-deoxy-glucose-positron emission tomography scan in non-small-cell lung cancer: an analysis of 125 cases. *J Clin Oncol* 1999; 17: 3201-6.
32. Tsutani Y, Miyata Y, Misumi K et al. Difference in prognostic significance of maximum standardized uptake value on ^{18}F-fluoro-2-deoxyglucose positron emission tomography between adenocarcinoma and squamous cell carcinoma of the lung. *Jpn J Clin Oncol* 2011; 41: 890-6.
33. Huang B, Chan T, Kwong DL et al. Nasopharyngeal carcinoma: investigation of intratumoral heterogeneity with ^{18}F-FDG PET/CT. *Am J Roentgenol* 2012; 199: 169-74.
34. Tixier F, Le Rest CC, Hatt M et al. Intratumor heterogeneity characterized by textural features on baseline ^{18}F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer. *J Nucl Med* 2011; 52: 369-78.

35. Kim YI, Kim YJ, Paeng JC et al. Heterogeneity index evaluated by slope of linear regression on ^{18}F-FDG PET/CT as a prognostic marker for predicting tumor recurrence in pancreatic ductal adenocarcinoma. *Eur J Nucl Med Mol Imaging* 2017; 44: 1995-2003.
36. Zhang H, Wroblewski K et al. Prognostic value of metabolic tumor burden from ^{18}F-FDG PET in surgical patients with non-small-cell lung cancer. *Acad Radiol* 2013; 20: 32-40.
37. Schmidt-Hansen M, Baldwin DR, Hasler E. PET/CT for assessing mediastinal lymph node involvement in patients with suspected resectable non-small cell lung cancer. *Cochrane Database Syst Rev* 2014; (11): CD009519.
38. Gerlinger M, Rowan AJ, Horswell S et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med* 2012; 366: 883-92.
39. Keyes JW Jr. SUV: standard uptake or silly useless value? *J Nucl Med* 1995; 36: 1836-9.