FACULDADE DE MEDICINA
UNIVERSIDADE D
COIMBRA

MESTRADO INTEGRADO EM MEDICINA – TRABALHO FINAL

ANA LUÍSA GOMES ROCHA

Metabolic active tumor volume quantified on [^{18}F]FDG PET/CT further stratifies TNM stage IV non-small cell lung cancer patients

ARTIGO CIENTÍFICO ORIGINAL

OUTRAS ÁREAS DA MEDICINA

Trabalho realizado sob a orientação de:

DOUTORA PAULA ALEXANDRA AMADO LAPA

PROFESSORA BÁRBARA CECÍLIA BESSA DOS SANTOS OLIVEIROS PAIVA

ABRIL/2021

ABRIL/2021

METABOLIC ACTIVE TUMOR VOLUME QUANTIFIED ON [^{18}F]FDG PET/CT FURTHER STRATIFIES TNM STAGE IV NON-SMALL CELL LUNG CANCER PATIENTS

O volume tumoral metabolicamente ativo do corpo inteiro quantificado na PET/CT com [^{18}F]FDG otimiza a estratificação dos doentes com cancro do pulmão de não pequenas células estadio TNM IV

Ana Luísa Gomes Rocha[1]; Bárbara Cecília Bessa Dos Santos Oliveiros Paiva[1], PhD; Paula Alexandra Amado Lapa[1,2], MD, PhD

1- Faculty of Medicine, University of Coimbra, Coimbra, Portugal

2- Department of Nuclear Medicine, Centro Hospitalar e Universitário de Coimbra, Coimbra, Portugal

Ana Luísa Gomes Rocha

Azinhaga de Santa Comba, Celas
3000-548 Coimbra

aluisagrocha@gmail.com

TABLE OF CONTENTS

ABBREVIATIONS..... 5

ABSTRACT 7

RESUMO 8

BACKGROUND..... 9

METHODS 11

Study design 11

Study population 11

[^{18}F]FDG PET/CT acquisition protocol 11

Data collection.....12

Statistical analysis 12

RESULTS 14

Patient characteristics 14

MTV_{WB} as a predictor of overall survival 15

MTV_{WB} with a cut-off point as a predictor of overall survival in stage IV patients.....17

MTV_{WB} with a cut-off point as a predictor of overall survival in each subsample – stage IVA and IVB.....17

DISCUSSION.....19

CONCLUSION 21

REFERENCES 22

ANNEX I 26

ABBREVIATIONS

[^{18}F]FDG PET/CT: 2-deoxy-2-[18F]fluoro-D-glucose Positron Emission Tomography/ Computed Tomography

3-D: Three Dimensional

AJCC: American Joint Committee on Cancer

Ca.: Carcinoma

CHUC: Centro Hospitalar e Universitário de Coimbra

CI: Confidence Interval

C Index: Harrell's Concordance Index

cm: centimeter

cm^3: cubic centimeter

CT: Computed Tomography

cTNM staging: Clinical Tumor, Node, Metastasis staging

EMST: Estimated Mean Survival Time

FU: Follow-Up

GE: General Electric

HR: Hazard Ratio

IMB: International Business Machines Corporation

IQR: Interquartile Range

kg: kilogram

kV: kilovolt

mA: milliampere

Max: Maximum

MBq: megabecquerel

MBq.min.bed^{-1}kg^{-1}: megabecquerel.minute.bed^{-1}kilogram^{-1}

Min: Minimum

min.bed^{-1}: minute.bed^{-1}

mm: millimeter

MTV: Metabolic active Tumor Volume

MTV_{WB}: Metabolic active Tumor Volume of the Whole Body

n: number of individuals

NSCLC: Non-Small Cell Lung Carcinoma

NY: New York

OS: Overall Survival

***p*:** *p-value*

PET: Positron Emission Tomography

PET/CT: Positron Emission Tomography/Computed Tomography

PET_VCAR: Positron Emission Tomography with Volume Computer Assisted Reading

SD: Standard Deviation

SE: Standard Error

SPSS: Statistical Package for the Social Sciences

SUV: Standardized Uptake Value

SUV_{max}: Maximum Standardized Uptake Value

SUV_{mean}: Mean Standardized Uptake Value

TLG: Total Lesion Glycolysis

TLG_{WB}: Total Lesion Glycolysis of the Whole Body

TNM staging: Tumor, Node, Metastasis staging

USA: United States of America

VUE: Virtually Unenhanced

vxtl: verxatile

WI: Wisconsin

ABSTRACT

Introduction: Stage IV non-small cell lung cancer (NSCLC) is a highly aggressive cancer with low survival rates. TNM staging has been recently revised to accommodate more substages, specifically substages IVA and IVB, for better stratification of patients. However, this population is rather heterogeneous and could benefit from more comprehensive staging algorithms. [^{18}F]FDG PET/CT provides volumetric parameters such as metabolic active tumor volume of the whole body (MTV_{WB}), which reflects tumor burden. The primary aim of this study was to assess whether MTV_{WB}, quantified in initial staging [^{18}F]FDG PET/CT, could further stratify stage IV patients, over standard cTNM staging.

Methods: A group of 160 patients submitted to initial staging [^{18}F]FDG PET/CT, between July 2010 and May 2020, and diagnosed with stage IV NSCLC, were retrospectively evaluated. MTV_{WB} was quantified and cTNM staging was recorded. Univariate and multivariate Cox regressions were carried out to assess correlation with overall survival (OS). C-statistic was used to test predictive power, and Kaplan-Meier survival curves with Log-Rank tests were performed to compute statistical differences between strata from dichotomized variables and to calculate the estimated mean survival time (EMST). Survival rates at one and five years were calculated.

Results: There were 70 (43.8%) stage IVA patients and 90 (56.3%) stage IVB patients. MTV_{WB} was a statistically significant predictor of OS on univariate (p<0.0001) and multivariate analyses (p<0.0001). A multivariate model with MTV_{WB} (C index±SE=0.657±0.024) was a significantly better predictor than the one with cTNM (C index±SE=0.544±0.028) (p=0.003). An EMST of 29.207±3.627 months (95% CI: 22.099-36.316) and EMST of 10.904±1.171 months (95% CI: 8.609-13.199) (Log-Rank: p<0.0001) were determined, respectively, for patients with MTV_{WB}<104.3 and MTV_{WB}≥104.3. In subsamples of stage IVA (cut-off point=114.5) and IVB patients (cut-off point=191.1), statistically significant differences between EMST were also reported, with *p-values* of 0.0001 and 0.0002, respectively. In both substages and in the entire cohort, patients with MTV_{WB} ≥cut-off points had lower EMST and lower survival rates.

Conclusion: The baseline metabolic active tumor volume of the whole body, measured on [^{18}F]FDG PET/CT for staging purposes, further stratifies stage IV NSCLC patients. This parameter is an independent predictor of overall survival and provides valuable prognostic information over cTNM staging.

Keywords: Non-Small Cell Lung Cancer, TNM Staging, PET-CT, Tumor Burden, Prognostic Factor

RESUMO

Introdução: O cancro do pulmão de não pequenas células (*NSCLC*) estadio IV é uma neoplasia agressiva e com taxas de sobrevivência baixas. O estadiamento TNM foi revisto recentemente, de forma a incluir dois novos subestadios, IVA e IVB, com o objetivo de estratificar melhor os doentes. No entanto, esta população é bastante heterogénea e beneficiaria de algoritmos de estadiamento mais abrangentes. A PET/CT com [^{18}F]FDG providencia parâmetros volumétricos como o volume tumoral metabolicamente ativo do corpo inteiro (MTV_{WB}), que reflete a carga tumoral. Assim, o objetivo deste estudo foi compreender se o MTV_{WB}, quantificado na PET/CT com [^{18}F]FDG de estadiamento, conseguiria otimizar a estratificação dos doentes estadio IV, para além do estadiamento cTNM.

Métodos: Um grupo de 160 doentes submetidos a PET/CT com [^{18}F]FDG de estadiamento, desde julho de 2010 a maio de 2020, e diagnosticados com *NSCLC* estadio IV, foram avaliados retrospetivamente. Foi quantificado o MTV_{WB} e registado o estadiamento cTNM. Realizaram-se regressões de Cox univariadas e multivariadas para avaliar a correlação com a sobrevivência global (*OS*). Utilizou-se *C-statistic* para testar poder preditivo, e curvas *Kaplan-Meier* com testes *Log-Rank* para estabelecer diferenças entre estratos provenientes de variáveis dicotomizadas, assim como para calcular tempos de sobrevivência média estimada (*EMST*). Foram calculadas as taxas de sobrevivência a 1 e 5 anos.

Resultados: Registaram-se 70 (43,8%) doentes com estadio IVA e 90 (56,3%) com IVB. Este parâmetro foi um preditor significativo nas análises univariada (p<0,0001) e multivariada (p<0,0001). Um modelo multivariado com MTV_{WB} (*C index*±EP=0,657±0,024) foi melhor preditor que um modelo multivariado com cTNM (*C index*±EP=0,544±0,028) (p=0,003). O *EMST* dos doentes com MTV_{WB} <104,3 foi de 29,207±3,627 meses (95% IC: 22,099-36,316), enquanto doentes com MTV_{WB} ≥104,3 tiveram um *EMST* de 10,904±1,171 meses (95% IC: 8,609-13,199) (*Log-Rank*: p<0,0001). Nas subamostras de doentes IVA (ponto de corte=114,5) e de doentes IVB (ponto de corte=191,1), também foram reportadas diferenças significativas entre *EMST*, com valores p de 0,0001 e 0,0002, respetivamente. Nos dois subestadios e na amostra total, os doentes com MTV_{WB} ≥pontos de corte tiveram *EMST* e taxas de sobrevivência inferiores.

Conclusão: O MTV_{WB}, quantificado na PET/CT com [^{18}F]FDG de estadiamento, otimiza a estratificação dos doentes com *NSCLC* estadio IV. Este parâmetro é um preditor independente de *OS* e providencia informação prognóstica, adicional ao estadiamento cTNM.
Palavras chave: Cancro do Pulmão de não Pequenas Células, Classificação TNM, PET CT, Carga Tumoral, Fatores Prognósticos

BACKGROUND

Lung cancer is the main culprit of cancer-related mortality worldwide and constitutes the type of cancer most commonly diagnosed in both genders combined. [1] These tumors can be broadly divided into two main histological categories: small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC), and the latter encompasses 85% of all lung cancer cases. [2] For optimal management of NSCLC, it is crucial to classify patients with tumor, node and metastasis (TNM) staging. These tumors are frequently not diagnosed until metastatic disease is present [2,3] and the five-year overall survival (OS) rates are still extremely poor in stage IV patients. [4]

Comprehensive staging remains the most important tool for prognostic purposes. [5] However, it may not always provide the most reliable prediction since each stage comprises a highly heterogeneous population. For instance, the novel 8th edition of TNM has recently segregated stage IV into two separate categories – IVA and IVB – since the prior single stage IV category did not account for the prognostic differences in these two new substages. [6] Therefore, prognostic factors besides the TNM staging system must be studied in order to better stratify patients and improve decision-making when selecting risk-adapted therapies. Other widely known patient-specific prognostic markers are age, gender, performance status and weight loss. [7]

Currently, [^{18}F]FDG PET/CT represents a key component of the diagnostic algorithm of NSCLC as well as of the evaluation of response to therapy and detection of recurrent disease. [8-10] Maximum standardized uptake value (SUV_{max}), which is defined as the value of the voxel showing the highest uptake, is the main parameter used in clinical practice to measure [^{18}F]FDG uptake. [8] Moreover, there have been several successful studies which demonstrated that SUV_{max} yields prognostic information. [9,11] Despite that, some concerns arise when measuring SUV_{max}, such as random noise causing a single 'hot' pixel rather than an accurate altered uptake in the body. [12]

Furthermore, since SUV_{max} only constitutes a semi-quantitative measure, quantitative parameters including metabolic active tumor volume (MTV) and total lesion glycolysis (TLG) have been assessed to accurately reflect tumor burden. MTV consists in the metabolic active tumor volume measured on PET/CT with a segmentation technique, whilst TLG is calculated by multiplying MTV by SUV_{mean}. Similarly to SUV_{max}, these volumetric parameters have been extensively studied for their ability to predict disease progression in NSCLC [12-15] and in other types of cancer. [16-18] Lee *et al.* were the first to demonstrate that MTV was a prognostic

factor independent from other established markers in lung cancer. [19] They also hypothesized that this result implied that some prognostic markers, mainly stage, could simply depend on more significant underlying factors such as tumor burden. [19] Some studies even postulated that MTV was superior to SUV_{max} [13,20] in terms of prognostic value and, more importantly, not inferior to TNM staging itself. [21]

Precisely, our center has previously conducted a study regarding this matter, in which the prognostic significance of MTV of the whole body (MTV_{WB}) was compared to the stratifying power of cTNM staging in a cohort of all stages. [21] In fact, this previous work proved that MTV_{WB} further stratified NSCLC patients, and proposed a new index containing both cTNM staging and MTV_{WB}. However, these results were achieved before the new division of stage IV, thus we found it relevant to verify whether MTV_{WB} would still carry prognostic importance in this group of patients, even considering this separation.

In order to expand on the aforementioned hypothesis, the primary aim of this study was to demonstrate whether MTV_{WB} could further stratify stage IV NSCLC patients, over the standard cTNM staging system, considering that mortality of this stage remains so high and optimal prognostic algorithms are needed to enroll these patients in certain clinical trials. This was achieved by testing its stratifying power and comparing its OS predictive ability with that of conventional cTNM staging.

METHODS

Study design

This retrospective study was conducted in the Department of Nuclear Medicine of the Centro Hospitalar e Universitário de Coimbra (CHUC) in February 2021. Ethical approval for the study was obtained from the CHUC Ethics Committee and principles from the Declaration of Helsinki were fully met. Informed consent was not required for this type of retrospective analysis.

Study population

We conducted a retrospective review of the medical records of patients diagnosed with NSCLC in our institution between 2010 and 2020. We identified the 160 consecutive patients based on the following inclusion criteria: 1) all patients had histological confirmation of the disease, 2) they were submitted to [^{18}F]FDG PET/CT at our institution for initial staging and 3) they were stage IV at diagnosis. These PET/CT scans were performed from July 2010 to May 2020. Exclusion criteria comprised: 1) presence of brain metastases (excluded by magnetic resonance) and 2) presence of other past or concurrent malignancies.

The PET/CT scans were performed before any therapeutic intervention. After attribution of cTNM staging and histological characterization of the lung tumor, patients were treated according to the most appropriate therapeutic strategies for their clinical condition, in accordance with the good practice guidelines at the time of treatment.

[^{18}F]FDG PET/CT acquisition protocol

This was a monocenter study and the [^{18}F]FDG PET/CT scans were conducted according to the institution's existing protocol: patients completed a 6-hour fast and their glycemic levels were below 144 mg/dL prior to intravenous [^{18}F]FDG administration. The administered activities were calculated based on the European Association of Nuclear Medicine Guidelines for tumor imaging on [^{18}F]FDG PET/CT – minimum [^{18}F]FDG (MBq) recommended for systems that apply a PET bed overlap of ≤30% = 14 (MBq.min.bed^{-1}kg^{-1}) × patient weight (kg)/emission acquisition duration per bed position (min.bed^{-1}). [22] Images were then acquired, also according to these guidelines, after the recommended 60-minute interval, with a range of 55 – 75 min. [22] The variations observed in the administered activities and biodistribution times were associated with the usual conditions of clinical practice. [23] Patients were positioned in dorsal decubitus and whole-body images were acquired using a General Electric Discovery ST PET/CT scanner (GE Healthcare, Waukesha, WI, USA). The acquisition parameters of CT for attenuation correction and anatomic mapping were as follows: 120 kV, smart mA (with current values between 10 and 200 mA and noise index 35), pitch 1.5:1, rotation 0.5 seconds

and slice thickness 3.75 mm. The PET emission study was obtained in 3-D mode with an acquisition time of 3 minutes per table position, as per the manufacturer's recommendations. The collected data were reconstructed with a Field Of View diameter of 70 cm and 256 × 256 matrix using the VUE Point 3-D iterative reconstruction algorithm, with two iterations, 35 subsets and 4 mm full width at half maximum post-reconstruction filter.

Data collection

All data was transcribed in a randomized database sheet. Age at the time of PET/CT, gender and the cTNM stage assigned to each patient were recorded. In order to be consistent throughout our study, we reviewed the group of patients whose staging had been previously performed with the AJCC 7th edition of the TNM staging and grouped them into stage IVA and IVB according to the AJCC 8th edition guidelines. [6] Histological types were also recorded.

The [^{18}F]FDG PET/CT scans were retrospectively evaluated on a dedicated post-processing workstation (Advanced Windows 4.4 GE Medical Systems, Milwaukee, USA). Each patient's lesions were delineated and evaluated using the Volume Computer Assisted Reading (PET_VCAR) software (version vxtl_8_3_65). The PET_VCAR software generated whole body 3-D regions of interest, based on the pre-defined threshold SUV value of 2.5. Regions corresponding to physiological uptake and/or uptake in benign lesions were manually excluded based on consensus between two nuclear medicine specialists. After this initial post processing step, 3-D regions of interest, corresponding to the primary lung tumor and all metastatic lesions, were obtained. A quantitative analysis was performed by the software to calculate MTV_{WB}. Additionally, SUV_{max} was also logged.

The primary endpoint of the study was OS. OS was calculated from the date of the initial baseline staging PET/CT scan to the date of death from any cause, based on the follow-up and records described above. The patients last known to be alive were censored at the date of the end of the study (February 11, 2021).

Statistical analysis

The values of the quantitative data were presented with minimum-maximum (mean ± standard deviation) or median (interquartile range), categorical data with n (%), and OS times with the estimated mean.

Univariate analyses using Cox proportional hazards regression were run for the total study population to assess the relationship between all the variables logged and survival. MTV_{WB},

cTNM staging, SUV_{max} and the other patient specific factors such as age at the time of PET/CT, gender and histological type were also submitted to a multivariate Cox regression.

The OS time predictive abilities of cTMN staging and MTV_{WB} were evaluated using Harrell-Concordance indexes (higher values indicating better discriminatory power). Multivariate Cox regressions adjusted for age, gender and histological type were run for both cTNM staging and MTV_{WB} separately. C indexes of each multivariate model were then determined with R software package 'SurvComp', that also provided standard errors, confidence intervals and *p-values* for both. Afterwards, the OS predictive abilities of both multivariate models were compared using the same package. [24]

An optimal cut-off point for MTV_{WB} was computed for stage IV patients using the 'cutp' function of the R software package 'SurvMisc'. [25] The cut-off point with the lowest *p-value* was selected. The total study population was then divided into two groups based on the cut-off point chosen. Kaplan-Meier analysis with the Log-Rank test was used to compare estimated mean survival time between the two groups. The one-year and five-year survival rates were computed and compared between subjects above and below the cut-off point. Estimated mean survival times and survival rates at one and five years of stage IVA and IVB patients were also computed and compared for validation purposes.

The same procedures described above were used to select the best cut-off points for MTV_{WB} in each patient subgroup, defined by cTNM stages IVA and IVB, and to compare the estimated mean survival times, as well as the one-year and five-year survival rates, between subjects who were above and below the cut-off points.

A two-tailed *p-value* of less than 0.05 was considered statistically significant for all tests performed. Analyses were performed using SPSS software (version 27; Armonk, NY, USA: IBM Corp) and R software (R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

Patient characteristics

The descriptive analysis of age at the time of PET/CT, gender, histological type, SUV_{max} and MTV_{WB} in the 160 stage IV patients, divided in IVA and IVB, is shown in **Table 1**. There were 114 (71.3%) men and 46 (28.7%) women, aged 34 to 88 years (mean ± SD = 66.0 ± 10.829). Histological findings consisted mainly in adenocarcinoma (n = 102; 63.7%). The rest of the types are described thoroughly in **Table 1**. Stage-wise, 70 (43.8%) patients had stage IVA, whilst the other 90 (56.3%) patients had stage IVB NSCLC. Overall measurements of SUV_{max} ranged from 3.0-45.6 (mean ± SD = 14.6 ± 6.621). MTV_{WB} ranged from 0.2 to 1181.0 cm^3 (mean ± SD = 196.8 ± 234.419). Follow-up time ranged from 0.39 to 104.25 months (mean ± SD = 16.691 ± 17.840).

The distribution of MTV_{WB} values in the stage IVA and IVB subsamples is depicted in **Table 2**. The median value obtained from IVA patients (65.9) is considerably smaller than in stage IVB patients (168.6). Stage IVA patients had measurements from 0.2 to 1181.0 cm^3 (mean ± SD = 128.0 ± 191.9), whilst stage IVB patients ranged from 12.1 to 1093.2 cm^3 (mean ± SD = 250.3 ± 251.0).

Table 1: Characterization of the enrolled patients.

Characteristics	**Values**
Age at PET/CT (years)	Mean, 66.0; SD, 10.829; Range, 34-88
Gender	
- Male	114 (71.3%)
- Female	46 (28.7%)
Histological type	
- Adenocarcinoma	102 (63.7%)
- Epidermoid carcinoma	26 (16.3%)
- Adenosquamous carcinoma	11 (6.9%)
- Pleomorphic carcinoma	10 (6.3%)
- Sarcomatoid carcinoma	4 (2.5%)
- Adenomucinous carcinoma	7 (4.4%)
cTNM stage IV	
- IVA	70 (43.8%)
- IVB	90 (56.3%)
SUV_{max}	Mean, 14.6; SD, 6.621; Range, 3.0-45.6
MTV_{WB} (cm^3)	Mean, 196.8; SD, 234.419; Range, 0.2-1181.0
Follow-up time (months)	Mean, 16.691; SD, 17.840; Range, 0.39-104.25

cm^3, cubic centimeter; cTNM, clinical tumor, node, metastasis; MTV_{WB}, metabolic active tumor volume of the whole body; PET/CT, positron emission tomography/computed tomography; SD, standard deviation; SUV_{max}, maximum standardized uptake value.

Table 2: Descriptive analysis of MTV_{WB} in the subsamples stage IVA and IVB.

	Mean ± SD	Min	Max	Median	IQR
cTNM stage IVA	128.0±191.9	0.2	1181.0	65.9	24.1-141.9
cTNM stage IVB	250.3±251.0	12.1	1093.2	168.6	83.2-327.9

cTNM, clinical tumor, node, metastasis; IQR, interquartile range; Min, minimum; Max, maximum; MTV_{WB}, metabolic active tumor volume of the whole body; SD, standard deviation.

MTV_{WB} as a predictor of overall survival

In the univariate analyses, gender, cTNM stage, SUV_{max} and MTV_{WB} were statistically significant (**Table 3**). Age and histological type were not; however, we included them in the multivariate model since they are relevant baseline factors. In the multivariate analysis (**Table 3**), age remained significant (p=0.004), as well as stage (p=0.012). Notably, SUV_{max} was not an independent predictor of survival (p=0.256), contrary to MTV_{WB} which remained significant (p<0.0001) and proved to be an independent OS predictor.

Given these findings, predictive abilities of two multivariate models adjusted for gender, age and histological type, one containing cTNM and the other MTV_{WB}, were compared using C-statistic. The model with cTNM staging was not a statistically significant predictor (p=0.123), in contrast to the model containing MTV_{WB} (p<0.0001). The predictive ability of the model with MTV_{WB} was significantly better than the one with cTNM staging (p=0.003) (**Table 4**).

Table 3: Univariate and multivariate analyses. Association of overall survival with age, gender, histological type, cTNM stage, SUV_{max} and MTV_{WB}.

	Univariate models			Multivariate model		
	HR	**CI (95%)**	***p-value***	**HR**	**CI (95%)**	***p-value***
Age	1.014	0.997-1.031	0.104	1.029	1.009-1.049	**0.004**
Gender						
- Male			Reference			Reference
- Female	0.622	0.423-0.914	**0.016**	0.608	0.399-0.925	**0.020**
Histological type						
- Adenocarcinoma			Reference			Reference
- Epidermoid Ca.	1.115	0.708-1.754	0.639	1.021	0.633-1.647	0.931
- Adenosquamous Ca.	0.977	0.490-1.943	0.946	1.406	0.667-2.959	0.370
- Pleomorphic Ca.	2.622	1.301-5.286	**0.007**	1.183	0.547-2.561	0.668
- Sarcomatoid Ca.	4.097	1.462-11.474	**0.007**	3.244	1.078-9.757	**0.036**
- Adenomucinous Ca.	0.944	0.383-2.328	0.900	1.394	0.536-3.627	0.496
cTNM stage IV						
- IVA			Reference			Reference
- IVB	1.693	1.193-2.403	**0.0029**	1.622	1.111-2.369	**0.012**
SUV_{max}	1.026	1.005-1.048	**0.016**	1.015	0.989-1.042	0.256
MTV_{WB}	1.002	1.001-1.003	**<0.0001**	1.002	1.001-1.003	**<0.0001**

Statistically significant results are presented in bold.
Ca., carcinoma; CI, confidence interval; cTNM, clinical tumor, node, metastasis; HR, hazard ratio; MTV_{WB}, metabolic active tumor volume of the whole body; SUV_{max}, maximum standardized uptake value.

Table 4: Comparison of cTNM staging and MTV_{WB} overall survival predictive abilities.

	C index			Comparison
	C index ± SE	**CI (95%)**	***p-value***	***p-value***
Multivariate model with cTNM	0.544±0.028	0.488-0.599	0.123	**0.003**
Multivariate model with MTV_{WB}	0.657±0.024	0.609-0.704	**<0.0001**	

Statistically significant results are presented in bold.
CI, confidence interval; cTNM, clinical tumor, node, metastasis; MTV_{WB}, metabolic active tumor volume of the whole body; OS, overall survival; SE, standard error.

MTV_{WB} with a cut-off point as a predictor of overall survival in stage IV patients

In order to employ MTV_{WB} in clinical practice, we dichotomized the variable with a calculated cut-off point. The value of 104.3 (p<0.0001) was identified as the optimal cut-off point for the whole sample. Patients with MTV_{WB}<104.3 had an estimated mean survival time of 29.207 ± 3.627 months (95% CI: 22.099-36.316), while those with MTV_{WB}≥104.3 had an estimated mean survival time of 10.904 ± 1.171 months (95% CI: 8.609-13.199). There was a statistically significant difference in the estimated mean survival times, in months, (Log-Rank Chi-Square = 27.165; p<0.0001) between the two groups of patients (**Figure 1-a**).

The probability of survival above or below the MTV_{WB} cut-off point at one and five years after diagnosis was calculated. The one-year survival rate was 39 ± 6% (mean ± standard error) for patients with MTV_{WB}<104.3 and only 12 ± 4% for patients with MTV_{WB}≥104.3. The five-year survival rate was 9 ± 4% for patients with MTV_{WB}<104.3 and there were no survivors in the group of patients with MTV_{WB}≥104.3 (**Annex I - Table 1**).

MTV_{WB} with a cut-off point as a predictor of overall survival in each subsample – stage IVA and IVB

Although we proved that MTV_{WB} can further stratify stage IV patients, this may be evident considering that this stage has been recently separated into two substages. Therefore, it made sense to search for cut-off points in each subsample. Additionally, as expected, there was a statistically significant difference in estimated mean survival times between stage IVA and IVB patients (p=0.003). The survival curves for these substages are shown in **Figure 1-b.** The estimated mean survival times as a function of cTNM stage are depicted in **Table 5.**

The identified optimal MTV_{WB} cut-off points were 114.5 (p=0.02) for stage IVA and 191.1 (p=0.005) for stage IVB. The estimated mean survival times of stage IVA and IVB patients, as a function of the respective MTV_{WB} are depicted in **Table 5.** Patients with values above the cut-off point at each category had worse prognosis. There was a statistically significant difference in estimated mean survival times, in months, between patient groups of both subsamples. Survival curves are represented in **Figures 1-c and 1-d** (p=0.0001 for stage IVA and p=0.0002 for stage IVB).

The one-year and five-year survival rates for the groups above and below the MTV_{WB} cut-off points at each substage were also determined. Patients with MTV_{WB} values above the cut-off points had lower survival rates than patients with MTV_{WB} values below them (**Annex I – Table 1**).

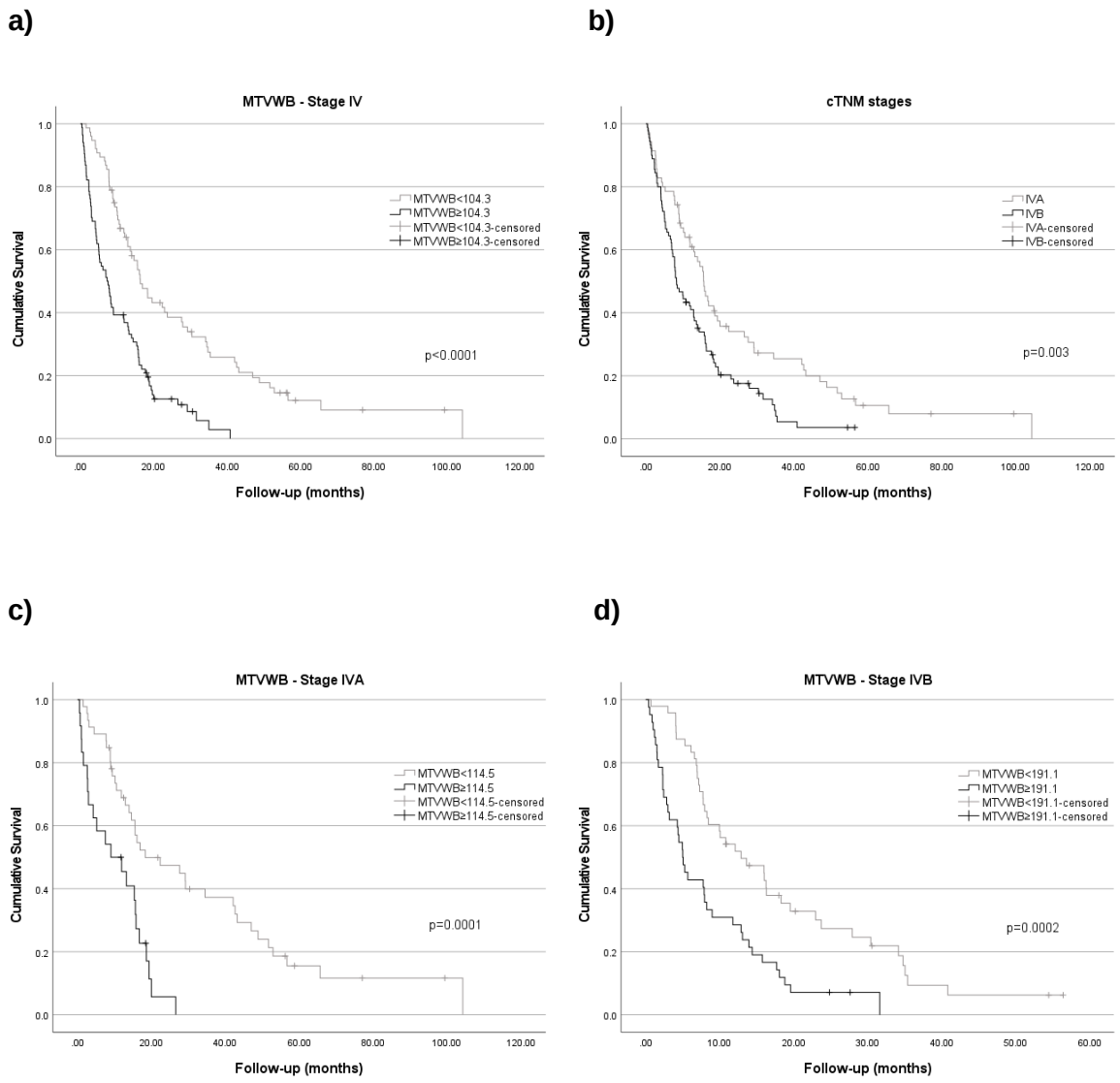

Figure 1: Kaplan-Meier curves comparing overall survival between groups as a function of: **a)** MTV_{WB} in the total study population (cTNM stage IV); **b)** cTNM stages – IVA and IVB in the total study population; **c)** MTV_{WB} in stage IVA patients; **d)** MTV_{WB} in stage IVB patients.

Table 5: Estimated mean survival time, in months, according to cTNM staging and according to the cut-off point defined for MTV_{WB} in each cTNM stage.

Stage	EMST±SE	CI (95%)	*p**	MTV_{WB}	EMST±SE	CI (95%)	*p**
IVA	26.374±3.657	19.205-33.542	**0.003**	<114.5	33.951±4.965	24.219-43.682	**0.0001**
				≥114.5	10.467±1.680	7.174-13.760	
IVB	13.833±1.464	10.964-16.702		<191.1	18.412±2.242	14.018-22.806	**0.0002**
				≥191.1	8.552±1.326	5.953-11.151	

Statistically significant results are presented in bold.
*Log-rank test.
CI, confidence interval; cTNM, clinical tumor, node, metastasis; EMST, estimated mean survival time; MTV_{WB}, metabolic active tumor volume of the whole body; SE, standard error.

DISCUSSION

A growing body of literature has been consistently proving that MTV adds stratification power. [12-14] However, this parameter has yet to be included in official staging guidelines. Thus, we found it relevant to further investigate this premise, particularly in patients with metastatic lesions, whose survival depends on the optimization of stratification and, consequently, the selection of risk-adapted therapies. In fact, our results have come to demonstrate that MTV_{WB} can subcategorize patients, specifically these patients with metastasized cancer. Firstly, MTV_{WB} was an independent predictor in multivariate analysis. Secondly, beyond the fact that MTV_{WB} proved to be a better OS predictor in comparison to cTNM staging through C-statistic, both survival curves as a function of an optimal MTV_{WB} cut-off point in stage IVA and IVB patients had statistically significant differences. Interestingly, it can also be noted that stage IVB patients below their computed cut-off point had a higher estimated mean survival time than stage IVA patients with values above their own cut-off point. We hypothesize that this could indicate that the updated cTNM is still not sufficient to classify these cases.

For validation purposes, we compared the five-year survival rates of our cohort to the ones documented in the proposals for new guidelines concerning the classification and cTNM staging of lung cancer. Our patients had a five-year survival rate of 7 ± 4% (mean ± standard error) in stage IVA and 0% in stage IVB. These guidelines show similar survival rates: 10% and 0%, respectively. [6] For this reason, we believe our sample was representative. Also, our cohort had similar numbers of stage IVA (70) and IVB (90) patients, allowing for a concise analysis of both groups.

Recently, Pu *et al*. conducted an analysis of a quite large sample of NSCLC patients of all stages, divided according to the new guidelines, where they validated a novel MTV_{WB} risk stratification system, [26] reporting results consistent with ours. Accordingly, Pellegrino and colleagues also concluded that MTV_{WB} was an independent predictor of OS in all stages. [27] On the other hand, they raised an important question about the absence of consensus on the optimal technique for MTV delineation. Our center used a threshold of 2.5 for SUV_{max}, in accordance with previous studies. [27,28] However, it should be noted that not all centers perform these measurements as described. We concur and corroborate that it is necessary to validate a certain method and threshold for more reproducible results. Of note, this choice of threshold intensity value did not affect the consistency of our measurements, since the same value was used for the whole sample. [19]

Moreover, there have been three previous studies regarding MTV and only stage IV patients, albeit considering the old cTNM staging guidelines. Firstly, a study with 92 consecutive patients with newly diagnosed stage IV NSCLC also indicated that MTV was a prognostic marker at the whole-body tumor burden level, and at the primary tumor level as well. [13] However, our sample was somewhat larger, and their results were obtained before the new guidelines had divided stage IV patients. This supports our hypothesis that this novel division is not sufficient, considering that our results were still concordant, and thus MTV_{WB} would improve stratification. In contrast, two other studies reported that only MTV from primary lung lesions at PET/CT for staging purposes had prognostic value. [29,30] Yet, Yoo *et al*. measured MTV of primary lung tumor and MTV-torso instead of MTV_{WB} in accordance with their country's standard protocol [29] and Lee and colleagues evaluated MTV at primary lesion level separately from MTV at the node and metastasis level; [30] therefore, comparisons cannot be accurately drawn.

Some of these studies assessed both MTV and the other volumetric value – TLG. Supposedly, TLG could be more promising since it combines volumetric and metabolic information. However, previous works did not report superiority of TLG, [29,31] and Zhang *et al*. added that, in future clinical practice, MTV_{WB} would be sufficient for measuring metabolic tumor burden in NSCLC, as there is no demonstrable advantage of TLG of the whole body (TLG_{WB}) over MTV_{WB}. [31] In any case, calculation of TLG is provided by the PET/CT station after MTV quantification; therefore, they could be easily logged together.

Additionally, our results substantiate that SUV_{max} may not be an independent predictor of OS, enhancing the importance of volumetric parameters. As Huang and colleagues posited, SUV_{max} may not be the most accurate predictor, since it only reflects a single-pixel value of maximal metabolic activity, that will show greater response to treatment, and, thus, less impact on outcome. [32] Conversely, MTV_{WB} will reflect metabolic changes throughout the entire tumor mass and, on top of that, it considers the whole-body tumor burden, yielding more precise information on prognosis.

It should be emphasized that therapeutic approaches have evolved drastically throughout this period of 10 years. Stage IV patients in 2010 were mostly treated with empirical cytotoxic therapies [33] whilst, more recently, immunotherapy and molecularly targeted therapies [33,34] were introduced, aiming to improve survival. This might have compromised our analysis since OS could have been considerably better in recent times even with larger tumor volumes, undermining the value of $MTV_{WB.}$ In spite of that, our results remained significant. Hence, we may conjecture that this marker is highly predictive, regardless of the selected treatment.

Despite the claimed advantages MTV_{WB} brings, its measurement can still be time-consuming. Our previous work presented a mean time of 5 minutes per patient, [21] yet this included early stage patients with low MTVs. In this study, we determined a mean time of roughly 25 minutes per case. However, with the advent of deep learning networks and computer-aided automatic processing, the need for handcrafted radiomic features of images will be eliminated and this process will become more efficient. [35,36]

Our study has some limitations. Firstly, we excluded patients with brain metastasis since [^{18}F]FDG PET/CT does not accurately characterize them. This might disregard a non-negligible number of patients, that can reach up to 26% of stage IV NSCLC cases. [37] Secondly, we were not able to retrieve the specific causes of death of each patient, considering some of them were unknown. However, the five-year survival rate of stage IV patients remains extremely low in both substages. [6] Hence, it is safe to assume that OS would be extremely close if not equal to disease-specific survival. In addition, the performance status at the time of PET/CT was also not documented since it was absent from some records. Nonetheless, poor performance status may simply depend on high tumor burden, [19] which is already assessed by MTV_{WB}.

Finally, the retrospective nature of this study encompasses already well-known drawbacks, and further prospective studies with larger cohorts should be performed to confirm our results. More importantly, multicenter projects should be carried out, in order to establish optimal MTV_{WB} cut-off points, suitable for the general NSCLC population. This would constitute a simple method of introducing this parameter into more comprehensive staging algorithms, allowing for improved planning of clinical trials and individualized therapeutic strategies.

CONCLUSION

The baseline metabolic active tumor volume of the whole body, measured on [^{18}F]FDG PET/CT for staging purposes, further stratifies stage IV NSCLC patients. This parameter is an independent predictor of overall survival and provides valuable prognostic information over cTNM staging. We suggest standardizing measurements between centers, as well as finding optimal cut-off points within stage IVA and IVB patients and incorporating them in official staging guidelines.

REFERENCES

1. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh JWW, Comber H, et al. Cancer incidence and mortality patterns in Europe: Estimates for 40 countries in 2012. European Journal of Cancer. 2013 Apr;49(6):1374–403.
2. Duma N, Santana-Davila R, Molina JR. Non–Small Cell Lung Cancer: Epidemiology, Screening, Diagnosis, and Treatment. Vol. 94, Mayo Clinic Proceedings. Elsevier Ltd; 2019. p. 1623–40.
3. Kocher F, Hilbe W, Seeber A, Pircher A, Schmid T, Greil R, et al. Longitudinal analysis of 2293 NSCLC patients: A comprehensive study from the TYROL registry. Lung Cancer. 2015;87(2):193–200.
4. William WN, Lin HY, Lee JJ, Lippman SM, Roth JA, Kim ES. Revisiting stage IIIB and IV non-small cell lung cancer: Analysis of the surveillance, epidemiology, and end results data. Chest. 2009 Sep 1;136(3):701–9.
5. Thakur MK, Gadgeel SM. Predictive and Prognostic Biomarkers in Non-Small Cell Lung Cancer. Seminars in Respiratory and Critical Care Medicine. 2016 Oct 1;37(5):760–70.
6. Goldstraw P. New Guidelines for the Classification and Staging of Lung Cancer: TNM Descriptor and Classification Changes in the 8th Edition.
7. Brundage MD, Davies D, Mackillop WJ. Prognostic factors in non-small cell lung cancer: A decade of progress. Vol. 122, Chest. 2002. p. 1037–57.
8. Groheux D, Quere G, Blanc E, Lemarignier C, Vercellino L, de Margerie-Mellon C, et al. FDG PET-CT for solitary pulmonary nodule and lung cancer: Literature review. Vol. 97, Diagnostic and Interventional Imaging. Elsevier Masson SAS; 2016. p. 1003–17.
9. Hanin FX, Lonneux M, Cornet J, Noirhomme P, Coulon C, Distexhe J, et al. Prognostic value of FDG uptake in early stage non-small cell lung cancer. European Journal of Cardio-thoracic Surgery. 2008 May;33(5):819–23.
10. Kandathil A, Kay FU, Butt YM, Wachsmann JW, Subramaniam RM. Role of FDG PET/CT in the eighth edition of TNM staging of non– Small cell lung cancer. Radiographics. 2018 Nov 1;38(7):2134–49.
11. Cerfolio RJ, Bryant AS, Ohja B, Bartolucci AA. The maximum standardized uptake values on positron emission tomography of a non-small cell lung cancer predict stage, recurrence, and survival. Journal of Thoracic and Cardiovascular Surgery. 2005;130(1):151–9.
12. Obara P, Pu Y. Prognostic value of metabolic tumor burden in lung cancer. Vol. 25, Chinese Journal of Cancer Research. Beijing Institute for Cancer Research; 2013. p. 615–22.

13. Liao S, Penney BC, Wroblewski K, Zhang H, Simon CA, Kampalath R, et al. Prognostic value of metabolic tumor burden on 18F-FDG PET in nonsurgical patients with non-small cell lung cancer. European Journal of Nuclear Medicine and Molecular Imaging. 2012 Jan;39(1):27–38.
14. Kim K, Kim SJ, Kim IJ, Seong Kim Y, Pak K, Kim H. Prognostic value of volumetric parameters measured by F-18 FDG PET/CT in surgically resected non-small-cell lung cancer. Vol. 33, Nuclear Medicine Communications. 2012. p. 613–20.
15. Chen HHW, Chiu NT, Su WC, Guo HR, Lee BF. Prognostic value of whole-body total lesion glycolysis at pretreatment FDG PET/CT in non-small cell lung cancer. Radiology. 2012 Aug;264(2):559–66.
16. La TH, Filion EJ, Turnbull BB, Chu JN, Lee P, Nguyen K, et al. Metabolic Tumor Volume Predicts for Recurrence and Death in Head-and-Neck Cancer. International Journal of Radiation Oncology Biology Physics. 2009 Aug 1;74(5):1335–41.
17. Berkowitz A, Basu S, Srinivas S, Sankaran S, Schuster S, Alavi A. Determination of whole-body metabolic burden as a quantitative measure of disease activity in lymphoma: a novel approach with fluorodeoxyglucose-PET. Vol. 29, Nuclear Medicine Communications. Wolters Kluwer Health | Lippincott Williams & Wilkins; 2008.
18. Roedl JB, Colen RR, Holalkere NS, Fischman AJ, Choi NC, Blake MA. Adenocarcinomas of the esophagus: Response to chemoradiotherapy is associated with decrease of metabolic tumor volume as measured on PET-CT. Comparison to histopathologic and clinical response evaluation. Radiotherapy and Oncology. 2008 Dec;89(3):278–86.
19. Lee P, Weerasuriya DK, Lavori PW, Quon A, Hara W, Maxim PG, et al. Metabolic Tumor Burden Predicts for Disease Progression and Death in Lung Cancer. International Journal of Radiation Oncology Biology Physics. 2007 Oct 1;69(2):328–33.
20. Liao S, Penney BC, Zhang H, Suzuki K, Pu Y. Prognostic Value of the Quantitative Metabolic Volumetric Measurement on 18F-FDG PET/CT in Stage IV Nonsurgical Small-cell Lung Cancer. Academic Radiology. 2012 Jan;19(1):69–77.
21. Lapa P, Oliveiros B, Marques M, Isidoro J, Alves FC, Nascimento Costa JM, et al. Metabolic tumor burden quantified on [18F]FDG PET/CT improves TNM staging of lung cancer patients. European Journal of Nuclear Medicine and Molecular Imaging. 2017;44(13):2169–78.
22. Boellaard R, Delgado-Bolton R, Oyen WJG, Giammarile F, Tatsch K, Eschner W, et al. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. Vol. 42, European Journal of Nuclear Medicine and Molecular Imaging. Springer Science and Business Media Deutschland GmbH; 2015. p. 328–54.

23. Graham MM, Badawi RD, Wahl RL. Variations in PET/CT methodology for oncologic imaging at U.S. academic medical centers: An imaging response assessment team survey. Journal of Nuclear Medicine. 2011 Feb 1;52(2):311–7.
24. Harrell FE, Lee KL, Mark DB. Tutorial In Biostatistics Multivariable Prognostic Models: Issues In Developing Models, Evaluating Assumptions And Adequacy, And Measuring And Reducing Errors. Vol. 15, Statistics In Medicine. 1996.
25. Contal C. An application of changepoint methods in studying the effect of age on survival in breast cancer. Vol. 30, Computational Statistics & Data Analysis. 1999.
26. Pu Y, Zhang JX, Liu H, Appelbaum D, Meng J, Penney BC. Developing and validating a novel metabolic tumor volume risk stratification system for supplementing non-small cell lung cancer staging. European Journal of Nuclear Medicine and Molecular Imaging. 2018 Nov 1;45(12):2079–92.
27. Pellegrino S, Fonti R, Mazziotti E, Piccin L, Mozzillo E, Damiano V, et al. Total metabolic tumor volume by 18F-FDG PET/CT for the prediction of outcome in patients with non-small cell lung cancer. Annals of Nuclear Medicine. 2019 Dec 1;33(12):937–44.
28. Im HJ, Pak K, Cheon GJ, Kang KW, Kim SJ, Kim IJ, et al. Prognostic value of volumetric parameters of 18F-FDG PET in non-small-cell lung cancer: a meta-analysis. European Journal of Nuclear Medicine and Molecular Imaging. 2015 Feb 1;42(2):241–51.
29. Yoo SW, Kim J, Chong A, Kwon SY, Min JJ, Song HC, et al. Metabolic tumor volume measured by F-18 FDG PET/CT can further stratify the prognosis of patients with stage IV non-small cell lung cancer. Nuclear Medicine and Molecular Imaging. 2012 Dec 1;46(4):286–93.
30. Lee JW, Lee SM, Yun M, Cho A. Prognostic value of volumetric parameters on staging and posttreatment FDG PET/CT in patients with stage IV non-small cell lung cancer. Clinical Nuclear Medicine. 2016;41(5):347–53.
31. Zhang H, Wroblewski K, Appelbaum D, Pu Y. Independent prognostic value of whole-body metabolic tumor burden from FDG-PET in non-small cell lung cancer. International Journal of Computer Assisted Radiology and Surgery. 2013;8(2):181–91.
32. Huang W, Fan M, Liu B, Fu Z, Zhou T, Zhang Z, et al. Value of metabolic tumor volume on repeated 18F-FDG PET/CT for early prediction of survival in locally advanced non-small cell lung cancer treated with concurrent chemoradiotherapy. Journal of Nuclear Medicine. 2014 Oct 1;55(10):1584–90.
33. Herbst RS, Morgensztern D, Boshoff C. The biology and management of non-small cell lung cancer. Vol. 553, Nature. Nature Publishing Group; 2018. p. 446–54.

34. Arbour KC, Riely GJ. Systemic therapy for locally advanced and metastatic non-small cell lung cancer: A review. JAMA - Journal of the American Medical Association. 2019 Aug 27;322(8):764–74.
35. Avanzo M, Stancanello J, Pirrone G, Sartor G. Radiomics and deep learning in lung cancer. Vol. 196, Strahlentherapie und Onkologie. Springer Science and Business Media Deutschland GmbH; 2020. p. 879–87.
36. Zhong Z, Kim Y, Zhou L, Plichta K, Allen B, Buatti J, Wu X. 3D Fully Convolutional Networks For Co-Segmentation Of Tumors On PET-CT Images. Proc IEEE Int Symp Biomed Imaging. 2018 Apr; 2018:228-231.
37. Waqar SN, Samson PP, Robinson CG, Bradley J, Devarakonda S, Du L, et al. Non–small-cell Lung Cancer With Brain Metastasis at Presentation. Clinical Lung Cancer. 2018 Jul 1;19(4):e373–9.

ANNEX I, Survival rates (%) of the entire cohort, divided by optimal cut-off points for the whole sample and subsamples, and divided by cTNM staging.

Annex I, Table 1: Survival rate (%) (mean ± standard error) according to the cut-off point defined for MTV_{WB} at each cTNM stage.

FU	IV		IVA		IVB	
	<104.3	≥104.3	<114.5	≥114.5	<191.1	≥191.1
1 year	**39±6**	**12±4**	**45±8**	**15±7**	**27±7**	**7±4**
5 year	9±4	0	11±5	0	0	0

Statistically significant results are presented in bold.
cTNM, clinical tumor, node, metastasis; FU, follow-up; MTV_{WB}, metabolic active tumor volume of the whole body.

Annex I, Table 2: Survival rate (%) (mean ± standard error) according to cTNM staging (IVA and IVB).

FU	Stage IV	
	IVA	IVB
1 year	**34±6**	**17±4**
5 year	7±4	0

Statistically significant results are presented in bold.
cTNM, clinical tumor, node, metastasis; FU, follow-up.