

[18F]FDG PET/CT-based response assessment of stage IV non-small cell lung cancer treated with paclitaxel-carboplatin-bevacizumab with or without nitroglycerin patches

Evelyn E. C. de Jong¹ · Wouter van Elmpt¹ · Ralph T. H. Leijenaar¹ · Otto S. Hoekstra² · Harry J. M. Groen³ · Egbert F. Smit^{4,5} · Ronald Boellaard⁶ · Vincent van der Noort⁷ · Esther G. C. Troost^{1,8,9} · Philippe Lambin¹ · Anne-Marie C. Dingemans¹⁰

Received: 27 May 2016 / Accepted: 16 August 2016 / Published online: 6 September 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract

Purpose Nitroglycerin (NTG) is a vasodilating drug, which increases tumor blood flow and consequently decreases hypoxia. Therefore, changes in [18F] fluorodeoxyglucose positron emission tomography ([18F]FDG PET) uptake pattern may occur. In this analysis, we investigated the feasibility of [18F]FDG PET for response assessment to paclitaxel-carboplatin-bevacizumab (PCB) treatment with and without NTG patches. And we compared the [18F]FDG PET response assessment to RECIST response assessment and survival.

Methods A total of 223 stage IV non-small cell lung cancer (NSCLC) patients were included in a phase II study (NCT01171170) randomizing between PCB treatment with or without NTG patches. For 60 participating patients, a baseline and a second [18F]FDG PET/computed tomography (CT) scan, performed between day 22 and 24 after the start of treatment, were available. Tumor response was defined as a 30 % decrease in CT and PET parameters, and was compared to

RECIST response at week 6. The predictive value of these assessments for progression free survival (PFS) and overall survival (OS) was assessed with and without NTG.

Results A 30 % decrease in SUVpeak assessment identified more patients as responders compared to a 30 % decrease in CT diameter assessment (73 % vs. 18 %), however, this was not correlated to OS (SUVpeak30 $p = 0.833$; CTdiameter30 $p = 0.557$). Changes in PET parameters between the baseline and the second scan were not significantly different for the NTG group compared to the control group (p value range 0.159–0.634). The CT-based (part of the [18F]FDG PET/CT) parameters showed a significant difference between the baseline and the second scan for the NTG group compared to the control group (CT diameter decrease of 7 ± 23 % vs. 19 ± 14 %, $p = 0.016$, respectively).

Conclusions The decrease in tumoral FDG uptake in advanced NSCLC patients treated with chemotherapy with and

✉ Evelyn E. C. de Jong
evelyn.dejong@maastro.nl

¹ Department of Radiation Oncology (MAASTRO), GROW-School for Oncology and Developmental Biology, Maastricht University Medical Centre, Maastricht, Netherlands
² Department of Nuclear Medicine & PET Research, VU University Medical Center, Amsterdam, Netherlands
³ Department of Pulmonary Diseases, University of Groningen and University Medical Center Groningen, Groningen, Netherlands
⁴ Department of Pulmonary Diseases, VU University Medical Center, Amsterdam, Netherlands
⁵ Department of Thoracic Oncology, The Netherlands Cancer Institute-Antoni van Leeuwenhoek Hospital, Amsterdam, Netherlands

⁶ Department of Nuclear Medicine and Molecular Imaging, University Medical Center Groningen, Groningen, Netherlands

⁷ Department of Biometrics, The Netherlands Cancer Institute-Antoni van Leeuwenhoek Hospital, Amsterdam, Netherlands

⁸ Institute of Radiooncology, Helmholtz-Zentrum Dresden-Rossendorf, Dresden, Germany

⁹ Department of Radiotherapy and Radiation Oncology, Medical Faculty and University Hospital Carl Gustav Carus of Technische Universität Dresden, Dresden, Germany

¹⁰ Department of Pulmonology, GROW-School for Oncology and Developmental Biology, Maastricht University Medical Centre, Maastricht, Netherlands

without NTG did not differ between both treatment arms. Early PET-based response assessment showed more tumor responders than CT-based response assessment (part of the [18F]FDG PET/CT); this was not correlated to survival. This might be due to timing of the [18F]FDG PET shortly after the bevacizumab infusion.

Keywords [18F]FDG PET/CT · Response assessment · Nitroglycerin · Bevacizumab · Stage IV NSCLC

Introduction

Molecular imaging with [18F] fluorodeoxyglucose positron emission tomography ([18F]FDG PET) has an established role in the staging of patients with non-small cell lung cancer (NSCLC). In addition, an increasing number of studies have shown that [18F]FDG PET is useful for early response assessment in NSCLC patients treated with cytotoxic chemotherapy [1–4].

Tumor hypoxia is a common phenomenon in lung cancer and it is related to poor prognosis due to treatment resistance [5–11]. Preclinical studies have shown that nitric oxide (NO)-donating drugs increase blood flow and thereby decrease hypoxia [12]. Nitroglycerin (NTG), a vasodilator, is such a drug. By increasing the tumor blood flow, NTG consequently augments antitumor drug delivery and inhibits hypoxia inducible factor (HIF-1 α) [13]. In preclinical models, administration of low doses of NTG, at least partially, reverses the hypoxia-induced resistance to anticancer drugs [14].

Yasuda et al. [15] showed that the combination of platinum-based chemotherapy and NTG improves overall survival (OS) in patients with stage IIIb/IV NSCLC. However, two recent randomized studies, including the Dutch NVALT12 study, could not confirm these results and no clinical effect was observed by the addition of NTG [16, 17].

Negative correlation between perfusion computed tomography (CT) and hypoxia PET on a population basis were also described in literature [18], suggesting that hypoxia is negatively correlated to tumor blood flow. Consequently, if treatment with NTG improved tumor perfusion, this could translate into a change in FDG uptake [13]. This concept was tested in the context of the randomized NVALT12 study that sought to investigate whether the addition of NTG to first-line paclitaxel-carboplatin-bevacizumab (PCB) chemotherapy would improve progression free survival (PFS).

In clinical practice, tumor response assessment is based on changes in tumor size, according to response evaluation criteria in solid tumors (RECIST) at week 6 [19, 20]. However, response monitoring is complex because the tumor has to change significantly in size and shape before a response is reliably detected by CT [21, 22]. This leads to an underestimation of the efficacy of cytostatic therapeutic agents that stabilize the disease, in contrast

to conventional cytotoxic drugs, which induce shrinkage of tumor dimensions in the case of tumor response [19]. Metabolic changes, measured by [18F]FDG PET, will occur earlier than changes in size and may, therefore, be used for early treatment response assessment. A decrease in metabolic activity of the primary tumor after one cycle of chemotherapy treatment is predictive for better outcome [1, 18, 23, 24].

In this paper, we investigated the feasibility of [18F]FDG PET for response assessment to PCB treatment with and without NTG patches. Furthermore, we compared the [18F]FDG PET response with both the commonly used RECIST and survival.

Materials and methods

Patient characteristics

In the multicentric NVALT12 trial (NCT01171170), 223 patients with metastatic non-squamous NSCLC were randomized between PCB with or without NTG (see [17] for patient inclusion criteria and treatment specifications) with the primary endpoint of that trial being PFS. Response was assessed every two cycles by the local investigator according to RECIST 1.1 based on CT imaging [20]. In patients undergoing [18F]FDG PET/CT at baseline as part of the standard work-up (median number of days between baseline scan and start treatment 17 days; range 73 days before treatment to 1 scan performed 1 day after the start of treatment), the protocol pre-specified a second [18F]FDG PET/CT between day 22 and 24 (after second chemotherapy infusion and with NTG application for patients in the experimental group; Fig. 1). To include more patients (17) in the analysis presented here, scans acquired with a time interval between the first chemotherapy and the second [18F]FDG PET/CT scan less than 35 days were accepted. This study was approved by the medical ethical committee and all patients provided informed written consent prior to any study handling.

Scan protocol

Injected [18F]FDG activity depended on individual patient and scanner characteristics, following the Netherlands protocol for standardization of [18F]FDG whole-body PET studies in multi-center trials (NEDPAS) [25], which was the precursor of the EANM guidelines, and images were reconstructed to institutional standards. Typically, a low-dose CT scan as part of the [18F]FDG PET/CT was made, according to institutional standards, and used for attenuation correction. Due to variations between the institutes, for quality control purposes, a spherical volume of interest (VOI) with a diameter of 3 cm was delineated in the right lobe

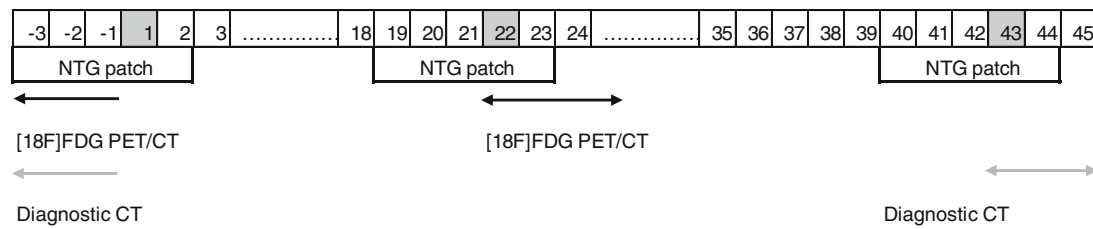


Fig. 1 NVALT12 trial timeline. At day one of the 21-day cycle, the paclitaxel-carboplatin-bevacizumab therapy is administered (*grey square*). The patients in the experimental arm wear the nitroglycerin (NTG) patch from day -3 to +2. The baseline [18F]FDG PET/CT is

performed before the start of chemotherapy and the second [18F]FDG PET/CT is performed between day 22 and 24 (*black arrow*). The baseline diagnostic CT is performed before the start of chemotherapy and repeated after every two cycles of chemotherapy (*grey arrow*)

of the liver [26]. This measurement was used as quality index and scans with a mean standardized uptake value (SUV) of the liver below 1.3 or above 3.0 were excluded from further analysis [27].

Early prediction of survival

The primary tumor was manually delineated by experienced radiation oncologists using a treatment planning system (Eclipse Version 11.0, Varian Medical Systems, Inc.) and used as the region of interest (ROI). A standard delineation protocol was used, which included fixed window/level settings of CT (lung: 1700/-300; mediastinum: 600/40). Patients without a measurable primary tumor on the baseline [18F]FDG PET/CT scan were excluded from analysis.

The maximum standardized uptake value (SUV_{max}), mean SUV (SUV_{mean}), peak SUV (SUV_{peak}; mean uptake in a sphere with a diameter of 1.2 cm [21]), total lesion glycolysis (TLG; TLG was defined by SUV_{mean} multiplied by the tumor volume), maximal CT diameter, and CT volume (number of voxels within the delineated ROI multiplied by the voxel size) were calculated in our institute on the [18F]FDG PET/CT scan (Matlab R2013a, The Mathworks, Natick, MA, USA) using an adapted version of CERR (Computational Environment for Radiotherapy Research) extended with in-house developed Radiomics image analysis software to extract imaging features [28, 29]. Early metabolic response was defined using relative changes in [18F]FDG PET uptake parameters of the primary tumor expressed as a percentage change from baseline. Patients were grouped according to a 30 % decrease in CT and PET parameters in the primary tumor ROI of the [18F]FDG PET/CT scan [26, 30, 31]. For the PET response assessment, SUV_{peak} was used and for the CT response assessment, CT diameter (CT was part of the [18F]FDG PET/CT) was used [26, 30]. The RECIST analysis performed during week 6 by the local investigators was used in the analysis to separate patients into responders and non-responders. The 30 % CT and

PET response assessments, performed after 3 weeks of therapy, were compared against the RECIST response assessment performed during week 6 by a specificity and sensitivity analysis.

Statistics

Since normality checks suggested an abnormal distribution for the changes in CT and PET parameters from baseline, non-parametric tests were used for the analysis of these variables. Comparison of the mean changes in CT and PET parameters from baseline for responders vs. non-responders was carried out by an independent samples Mann–Whitney *U* test. PFS was defined as the interval from randomization to progressive disease or death, whichever occurred first, and OS was defined as the interval from randomization to death from any cause. Differences in PFS and OS were investigated using Cox regression. For calculating the hazard ratio (HR), the different response assessment criteria were used, as a binary variable. To compare CT diameter and SUV_{max} response with survival, in the waterfall plots a survival cut-off of 6 months was used. This is the median PFS of the combined group (NTG group combined with control group). Statistical tests were based on a two-sided significance level, and the level of significance was set at 0.05. All statistics were performed in SPSS v.21 (IBM Corp. Released 2012, IBM SPSS Statistics for Windows, Version 21.0, Armonk, NY, USA).

Results

Patients

87 out of the 223 included patients in the randomized phase II study had two [18F]FDG PET/CT scans available with a measurable primary tumor; however, 27 patients were subsequently excluded for analysis due to poor image quality (see methods). Hence, 60 patients (characteristics in Table 1) had

Table 1 Patient characteristics

		Control	Experimental
Patients analyzed		31	29
Gender	Male	15	15
	Female	16	14
Age (mean, range)	[years]	59 (39–73)	59 (45–77)
WHO-PS	0	20 (65 %)	10 (34 %)
	1	10 (32 %)	18 (62 %)
	2	1 (3 %)	1 (3 %)
Smoker	Current	13 (42 %)	14 (48 %)
	Ex	14 (45 %)	12 (41 %)
	Never	4 (13 %)	3 (11 %)
Histology	Adeno	27 (86 %)	24 (83 %)
	Large cell	2 (7 %)	1 (3 %)
	Other	2 (7 %)	4 (14 %)
Survival (median, range)	PFS [months]	7 (3–25)	4 (1–11)
	OS [months]	13 (4–33)	9 (2–29)
RECIST response (week 6)	Complete response	0 (0 %)	0 (0 %)
	Partial response	9 (29 %)	5 (17 %)
	Stable disease	20 (65 %)	17 (59 %)
	Progressive disease	2 (6 %)	7 (24 %)
Baseline (mean, range)	CT diameter [cm]	6.8 (2.5–12.1)	6.7 (2.4–16.4)
	CT volume [cm ³]	101.5 (4.4–474.5)	89.8 (3.0–468.8)
	SUVmax	13.5 (3.4–28.9)	14.5 (3.6–44.6)
	SUVmean	5.7 (2.5–11.8)	6.2 (2.1–22.3)
	SUVpeak	11.0 (3.1–25.3)	12.1 (2.6–37.9)
	TLG [SUV*cm ³]	655.0 (23.4–4288.8)	638.4 (8.2–3467.2)

WHO-PS World Health Organization performance status, PFS progression-free survival, OS overall survival, RECIST response evaluation criteria in solid tumors, SUV standardized uptake value, TLG total lesion glycolysis

two evaluable consecutive [18F]FDG PET/CT scans (Fig. 2) with a median interval of 42 days. PFS and OS were similar for patients treated with PCB and PCB + NTG (Table 1).

Image characteristics

Experimental vs. control arm

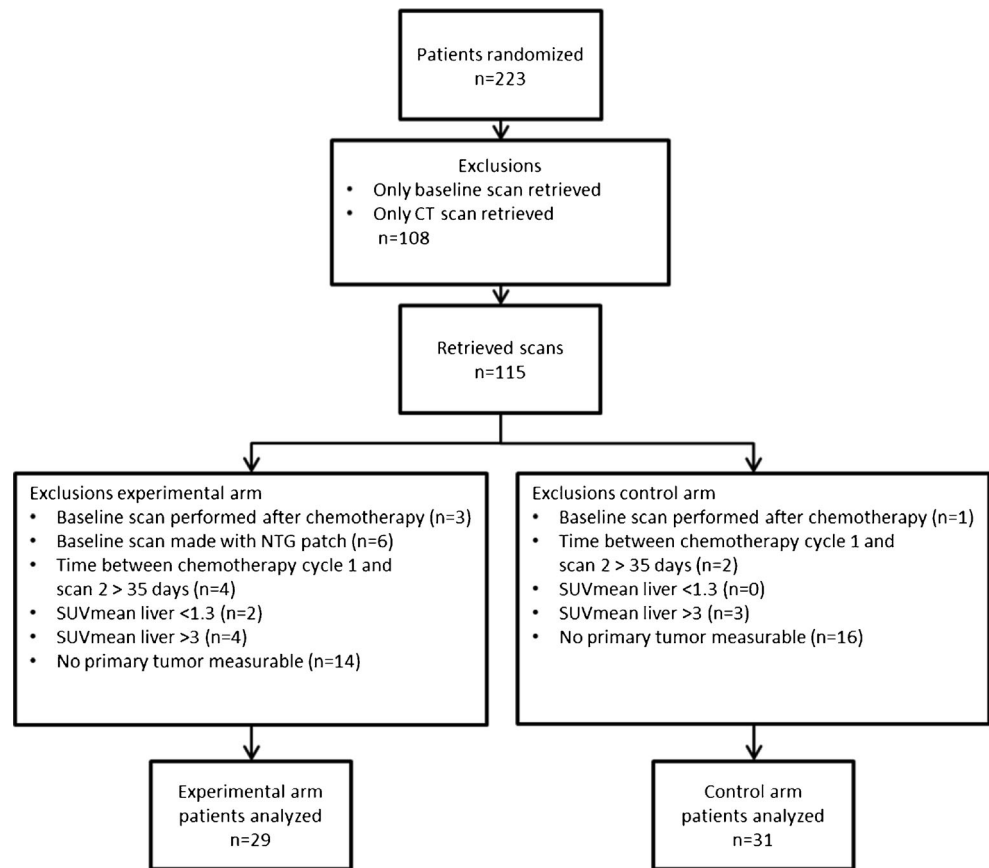
The mean decrease in SUVmax between the 31 patients treated with PCB (46 ± 27 %) and the 29 patients treated with PCB + NTG (42 ± 29 %) was not statistically significantly different ($p = 0.510$). The other PET parameters (SUVmean, SUVpeak and TLG) showed on average > 40 % decrease from baseline, but this was also not statistically significantly different between the experimental arm and the control arm (Fig. 3). Although for CT, part of the [18F]FDG PET/CT, in the control arm, the CT diameter decreased significantly more than in the experimental arm (19 ± 14 % vs. 7 ± 23 %; $p = 0.028$).

Early prediction of survival

According to the 30 % PET criteria, 74 % of patients in the control arm and 72 % of the patients in the experimental arm showed response after 3 weeks (median time interval 42 days). According to the 30 % CT criteria, 26 % of the patients in the control arm and 10 % of the patients in the experimental arm had a response. According to the RECIST analysis performed after 2 cycles (median time interval 56 days) by the local investigator, 29 % of the patients in the control arm had a response and 17 % of the patients in the experimental arm had a response (Table 1).

The predictive value of the 30 % CT-based and 30 % PET-based response assessments performed after 3 weeks (on the primary tumor) was assessed for response according to RECIST after 2 cycles (Table 2). The 30 % PET-based response assessment had a higher sensitivity compared to the 30 % CT-based response assessment but a lower specificity (Table 2).

Fig. 2 CONSORT diagram.
SUV: standardized uptake value



The 30 % CT-based and 30 % PET-based response assessments were for neither of the arms predictive for PFS nor OS (Table 3).

The changes in CT diameter and SUVmax between baseline and early response assessment were depicted in a waterfall plot showing that PET defined more patients as responders than CT (Fig. 4 and Table 2). However, this decline was not predictive for longer PFS (than 6 months).

Discussion

The hypothesis of the NVALT12 trial was that the addition of NTG, by increasing tumor blood flow and oxygenation status, would improve outcome. While the clinical study of the NVALT12 already showed that NTG did not improve outcome, in the current study, we investigated if we could predict outcome based on early response assessment using [18F]FDG PET imaging [17]. This image analysis study of the NVALT12 trial could not show a predictive value of [18F]FDG PET imaging for the evaluation of the addition of NTG to bevacizumab-containing chemotherapy when compared to control patients. In a previous study, the administration of NO donating drugs decreased hypoxia-induced resistance

to anticancer drugs in cancer cell lines [14]. In the NVALT12 trial, this could not be confirmed based on [18F]FDG PET analysis. This could be due to a lower NTG dose than that used in the Yasuda study [15], or to an interference with bevacizumab. From recent studies, it is known that FDG is only a moderate surrogate for hypoxia [32]. The study of Zegers et al. [33] showed that 42 ± 21 % of the primary tumor volume has a high FDG uptake ($SUV > 50$ % of SUV_{max}) of which 10 ± 12 % is hypoxic (high [18F]HX4 uptake $TBR > 1.4$), and that 3 % of the primary tumor volume outside the high FDG uptake volume is hypoxic as depicted by [18F]HX4 PET. In our study, we, therefore, only measure the effects of NTG on tumor metabolism and survival but not on hypoxia directly. Surprisingly, in nearly all patients, irrespective of treatment arm, a major decrease in FDG uptake was observed in the [18F]FDG PET scan performed after 3 weeks. Importantly, this [18F]FDG PET scan was acquired within 3 days after administration of the second cycle of chemotherapy. A study by van der Veldt et al. [34] showed that bevacizumab reduces tumor perfusion and [11C] docetaxel uptake in NSCLC, which was accompanied by rapid reduction in circulating levels of VEGF. This decrease in tumor blood flow after bevacizumab administration may explain the

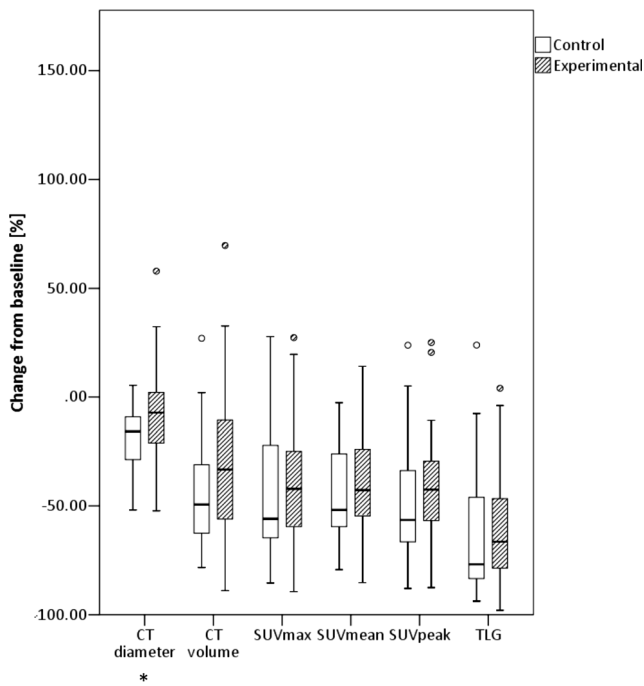


Fig. 3 Mean values and standard deviations for the CT- and PET-derived image parameters for the experimental arm and the control arm. *p* values of the independent samples Mann–Whitney *U* test of the mean change from baseline of the control arm vs. the mean change from baseline of the experimental arm (*significantly different for the experimental arm compared to the control arm with a significance level of 5 %). SUV: standardized uptake value; TLG: total lesion glycolysis

lower uptake of FDG in the tumor. Consequently, our results do not exclude the possibility that NTG decreases hypoxia.

A number of studies have demonstrated that changes in SUV parameters as early as the third week after the start of treatment are predictive for response to chemotherapy and PFS [1, 23, 24, 35]. A variety of approaches have been developed to measure the response, starting with

the World Health Organization (WHO) criteria and continuing to RECIST and RECIST 1.1 [20, 26, 36]. These criteria refer to an anatomical decrease in tumor diameter. However, this response must be viewed with some caution when one is trying to predict outcomes in therapies that may be more cytostatic than cytotoxic. With such therapies, lack of progression may be associated with a good improvement in outcome, even in the absence of major shrinkage of tumors [37]. Newer metrics such as PET may be more informative [38]. PET/CT-based response evaluation has proven to be valuable in chemotherapy [39]. Currently, two sets of treatment response criteria for PET are available: EORTC and PET response criteria in solid tumors (PERCIST) [30]. PERCIST operates with a fixed ROI of 1 cm³ in the most [18F]FDG-avid part of the single most metabolically active tumor in the patient at each PET/CT scan. In the current study, a specific ROI, defining the primary tumor, was used for response evaluation. A consideration for anatomic and functional imaging is that many of the changes in response are at the border zones between response groups.

These border zones are quite artificial, as changes in tumor size are on a continuous scale (Fig. 3). The comparison of 30 % CT-based and 30 % PET-based response assessment performed after 3 weeks (median time interval between scans 42 days) with the RECIST analysis performed in week 6 (median time interval between scans 56 days) showed that the RECIST analysis defined more patients as responders than the 30 % CT-based analysis performed after 3 weeks. This can be caused by the difference in timing but also due to the fact that for the 30 % CT-based analysis, only one lesion was measured while in RECIST, multiple lesions were measured. The 30 % PET-based response assessment performed after 3 weeks showed more responders than the RECIST analysis, which is probably caused by

Table 2 Comparison of 30 % CT-based and 30 % PET-based response assessment performed after 3 weeks with the RECIST response assessment of week 6

	RECIST Responder	RECIST Non-responder	Total
CT diameter decrease >30 %	9 (15 %)	2 (3 %)	11 (18 %)
CT diameter decrease <30 %	17 (28 %)	32 (54 %)	49 (82 %)
	26 (43 %)	34 (57 %)	60 (100 %)
	Sensitivity = 35 %	Specificity = 94 %	
SUVpeak decrease >30 %	23 (38 %)	21 (35 %)	44 (73 %)
SUVpeak decrease <30 %	3 (5 %)	13 (22 %)	16 (27 %)
	26 (43 %)	34 (57 %)	60 (100 %)
	Sensitivity = 88 %	Specificity = 38 %	

SUV standardized uptake value, RECIST response evaluation criteria in solid tumors

Table 3 The hazard ratios (HR) for 30 % PET- and CT-based response assessment with 95 % confidence interval and corresponding *p* values for OS and PFS are shown per parameter

SUV parameter		PFS		OS	
		HR (95 % CI)	<i>p</i> value	HR (95 % CI)	<i>p</i> value
30 % response assessment	SUVmax	1.048 (0.591–1.858)	0.874	1.025 (0.572–1.837)	0.934
	SUVmean	0.941 (0.527–1.680)	0.838	0.901 (0.501–1.619)	0.726
	SUVpeak	0.929 (0.514–1.679)	0.807	0.938 (0.515–1.706)	0.833
	TLG	0.706 (0.355–1.406)	0.323	1.511 (0.722–3.160)	0.273
	CTvolume	1.073 (0.617–1.866)	0.802	1.338 (0.740–2.419)	0.335
	CTdiameter	0.718 (0.370–1.390)	0.325	0.805 (0.390–1.662)	0.557

SUV standardized uptake value, TLG total lesion glycolysis, PFS progression free survival, OS overall survival, HR hazard ratio, CI confidence interval

decreased perfusion due to the bevacizumab treatment, which led to a decrease in FDG uptake for both treatment arms. The response assessment for PET was not influenced by NTG. A previous study of our group showed that after 3 weeks of treatment, five of nine patients were

classified as responder by CT while six of nine were classified as responders by [18F]FDG PET [40]. In the same study, patients with a metabolic response (decrease in SUV > 20 %) at week 3 had a longer PFS than those without (9.7 months vs. 2.8 months), while patients with a response on CT at week 3 did not have a significantly longer PFS than those without. These two findings combined showed that PET may be able to show treatment response earlier than CT. In the former study, [18F]FDG PET scans were performed before bevacizumab infusion, while in our study the [18F]FDG PET scan was performed shortly after bevacizumab infusion. This might have impacted the uptake of FDG. A study of Hoekstra et al. [41] also shows that [18F]FDG PET has additional value over conventional radiologic techniques for monitoring response in locally advanced NSCLC patients.

The scans used for this study were made within the scope of the Dutch multicenter NVALT12 phase II trial and 60/223 patients underwent 2 [18F]FDG PET scans with the second scan after cycle 2, but before day 35. For quality control purposes, only scans with a mean SUV in the liver between 1.3 and 3 were used, reducing the number of assessed patients in this analysis to only 60 of the 223 original patients. For the analysis, these 60 patients were also divided between the control and the experimental arm, which means the study cohort was limited in size, hampering in-depth subgroup analyses.

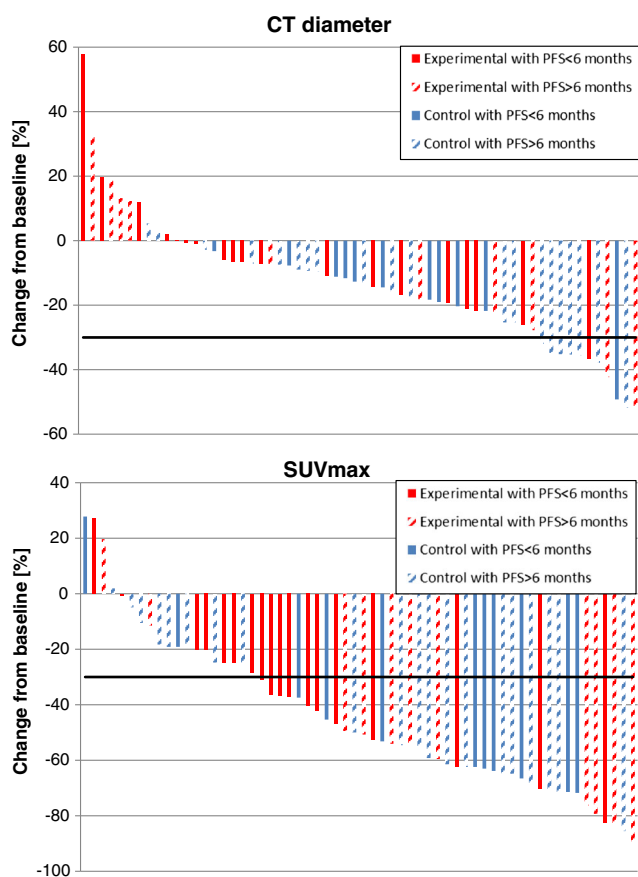


Fig. 4 Change in CT diameter (upper) and SUVmax (lower) from baseline in individual patients. Patients of the experimental arm are plotted in red, patients of the control arm in blue. The pattern-filled bars represent patients with a progression free survival longer than 6 months. The black line represents the used response threshold of 30 %. SUV: standardized uptake value; PFS: progression-free survival

Conclusion

The addition of NTG did not lead to enhanced reduction in FDG uptake compared to the control arm. Although PET-based response assessment identified more responders than CT-based response assessment, this did not correlate to progression-free survival or overall survival. This might be due to the timing of the [18F]FDG PET shortly after the bevacizumab infusion.

Compliance with ethical standards

Funding The NVALT12 is a multicentre randomized open-label parallel group phase II trial conducted by the Dutch Lung Physician Society (NVALT) and was supported by the Dutch Cancer Society (grant UM 2010–4883). This study is also funded by the Dutch Technology Foundation STW (grant n° 10696 DuCAT & n° P14-19 Radiomics STRaTegy), which is the applied science division of NWO and the Technology Programme of the Ministry of Economic Affairs.

Conflict of interest Author RL is employed by pfTheragnostics BV and is the Chief Scientific Officer of OncoRadiomics (a department of Health Innovations B.V.). Author AD is an advisory board member of Roche, Eli Lilly, Astra Zeneca, MSD and Pfizer. Author PL acknowledges financial support from ERC advanced grant (ERC-ADG-2015, n° 694812 - Hypoximmuno), the QuIC-ConCePT project, which is partly funded by EFPI A companies and the Innovative Medicine Initiative Joint Undertaking (IMI JU) under grant agreement no. 115151. Author PL also acknowledges financial support from the EU 7th Framework Program (ARTFORCE - n° 257144), Kankeronderzoekfonds Limburg from the Health Foundation Limburg and the Dutch Cancer Society. Authors EJ, WE, OH, HG, RB, VN and ET declare that they have no conflicts of interest.

Ethical approval All procedures performed in this study were in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki Declaration and its later amendments.

Informed consent Informed consent was obtained from all individual participants included in the study.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- de Geus-Oei LF, van der Heijden HF, Visser EP, Hermsen R, van Hoorn BA, Timmer-Bonte JN, et al. Chemotherapy response evaluation with 18F-FDG PET in patients with non-small cell lung cancer. *J Nucl Med Off Pub Soc Nuc Med.* 2007;48(10):1592–8. doi:10.2967/jnumed.107.043414.
- Vaidya M, Creach KM, Frye J, Dehdashti F, Bradley JD, El Naqa I. Combined PET/CT image characteristics for radiotherapy tumor response in lung cancer. *Radiother Oncol J Eur Soc Ther Radiol Oncol.* 2012;102(2):239–45. doi:10.1016/j.radonc.2011.10.014.
- Ordu C, Selcuk NA, Erdogan E, Angin G, Gural Z, Memis H, et al. Does early PET/CT assessment of response to chemotherapy predicts survival in patients with advanced stage non-small-cell lung cancer? *Medicine.* 2014;93(28), e299. doi:10.1097/MD.000000000000299.
- Grootjans W, de Geus-Oei LF, Troost EG, Visser EP, Oyen WJ, Bussink J. PET in the management of locally advanced and metastatic NSCLC. *Nat Rev Clin Oncol.* 2015;12(7):395–407. doi:10.1038/nrclinonc.2015.75.
- Hockel M, Vaupel P. Tumor hypoxia: definitions and current clinical, biologic, and molecular aspects. *J Natl Cancer Inst.* 2001;93(4):266–76.
- Harris AL. Hypoxia—a key regulatory factor in tumour growth. *Nat Rev Cancer.* 2002;2(1):38–47. doi:10.1038/nrc704.
- Verwer EE, Boellaard R, van der Veldt AA. Positron emission tomography to assess hypoxia and perfusion in lung cancer. *World J Clin Oncol.* 2014;5(5):824–44. doi:10.5306/wjco.v5.i5.824.
- Wouters BG, Wepler SA, Koritzinsky M, Landuyt W, Nuyts S, Theys J, et al. Hypoxia as a target for combined modality treatments. *Eur J Cancer.* 2002;38(2):240–57.
- Wouters BG, van den Beucken T, Magagnin MG, Lambin P, Koumenis C. Targeting hypoxia tolerance in cancer. *Drug Resist Updat Rev Comment Antimicrob Anticancer Chemother.* 2004;7(1):25–40. doi:10.1016/j.drug.2003.12.004.
- Dubois LJ, Niemans R, van Kuijk SJ, Panth KM, Parvathaneni NK, Peeters SG, et al. New ways to image and target tumour hypoxia and its molecular responses. *Radiother Oncol J Eur Soc Ther Radiol Oncol.* 2015;116(3):352–7. doi:10.1016/j.radonc.2015.08.022.
- Peeters SG, Zegers CM, Yaromina A, Van Elmpt W, Dubois L, Lambin P. Current preclinical and clinical applications of hypoxia PET imaging using 2-nitroimidazoles. *Q J Nucl Med Mol Imaging.* 2015;59(1):39–57.
- Jordan BF, Misson P, Demeure R, Baudelet C, Beghein N, Gallez B. Changes in tumor oxygenation/perfusion induced by the no donor, isosorbide dinitrate, in comparison with carbogen: monitoring by EPR and MRI. *Int J Radiat Oncol Biol Phys.* 2000;48(2):565–70.
- Yasuda H, Nakayama K, Watanabe M, Suzuki S, Fuji H, Okinaga S, et al. Nitroglycerin treatment may enhance chemosensitivity to docetaxel and carboplatin in patients with lung adenocarcinoma. *Clin Cancer Res Off J Am Assoc Cancer Res.* 2006;12(22):6748–57. doi:10.1158/1078-0432.CCR-06-1124.
- Matthews NE, Adams MA, Maxwell LR, Gofton TE, Graham CH. Nitric oxide-mediated regulation of chemosensitivity in cancer cells. *J Natl Cancer Inst.* 2001;93(24):1879–85.
- Yasuda H, Yamaya M, Nakayama K,asaki T, Ebihara S, Kanda A, et al. Randomized phase II trial comparing nitroglycerin plus vinorelbine and cisplatin with vinorelbine and cisplatin alone in previously untreated stage IIIB/IV non-small-cell lung cancer. *J Clin Oncol Off J Am Soc Clin Oncol.* 2006;24(4):688–94. doi:10.1200/JCO.2005.04.0436.
- Davidson A, Veillard AS, Tognela A, Chan MM, Hughes BG, Boyer M, et al. A phase III randomized trial of adding topical nitroglycerin to first-line chemotherapy for advanced nonsmall-cell lung cancer: the Australasian lung cancer trials group NITRO trial. *Ann Oncol Off J Eur Soc Med Oncol /ESMO.* 2015;26(11):2280–6. doi:10.1093/annonc/mdv373.
- Dingemans AM, Groen HJ, Herder GJ, Stigt JA, Smit EF, Bahce I, et al. A randomized phase II study comparing paclitaxel-carboplatin-bevacizumab with or without nitroglycerin patches in patients with stage IV nonsquamous nonsmall-cell lung cancer: NVALT12 (NCT01171170)dagger. *Ann Oncol Off J Eur Soc Med Oncol/ESMO.* 2015;26(11):2286–93. doi:10.1093/annonc/mdv370.
- van Elmpt W, Zegers CM, Reymen B, Even AJ, Dingemans AM, Oellers M, et al. Multiparametric imaging of patient and tumour heterogeneity in non-small-cell lung cancer: quantification of tumour hypoxia, metabolism and perfusion. *Eur J Nucl Med Mol Imaging.* 2016;43(2):240–8. doi:10.1007/s00259-015-3169-4.
- Hachemi M, Couturier O, Vervueren L, Fosse P, Lacoeyille F, Urban T, et al. [18F]FDG positron emission tomography within 2 weeks of starting erlotinib therapy can predict response in non-small cell lung cancer patients. *PLoS One.* 2014;9(2):e87629. doi:10.1371/journal.pone.0087629.

20. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer*. 2009;45(2):228–47. doi:10.1016/j.ejca.2008.10.026.
21. Baghaei H, Wong W-HG, Li H. Prin Pos Emiss Tomograp Imaging. 2013. doi:10.1007/978-1-4419-0802-5_1.
22. Erasmus JJ, Gladish GW, Broemeling L, Sabloff BS, Truong MT, Herbst RS, et al. Interobserver and intraobserver variability in measurement of non-small-cell carcinoma lung lesions: implications for assessment of tumor response. *J Clin Oncol Off J Am Soc Clin Oncol*. 2003;21(13):2574–82. doi:10.1200/jco.2003.01.144.
23. Kahraman D, Scheffler M, Zander T, Nogova L, Lammertsma AA, Boellaard R, et al. Quantitative analysis of response to treatment with erlotinib in advanced non-small cell lung cancer using 18F-FDG and 3'-deoxy-3'-18F-fluorothymidine PET. *J Nucl Med Off Pub Soc Nuc Med*. 2011;52(12):1871–7. doi:10.2967/jnumed.111.094458.
24. Nahmias C, Hanna WT, Wahl LM, Long MJ, Hubner KF, Townsend DW. Time course of early response to chemotherapy in non-small cell lung cancer patients with 18F-FDG PET/CT. *J Nucl Med Off Pub Soc Nuc Med*. 2007;48(5):744–51. doi:10.2967/jnumed.106.038513.
25. Boellaard R, Oyen WJ, Hoekstra CJ, Hoekstra OS, Visser EP, Willemsen AT, et al. The Netherlands protocol for standardisation and quantification of FDG whole body PET studies in multi-centre trials. *Eur J Nucl Med Mol Imaging*. 2008;35(12):2320–33. doi:10.1007/s00259-008-0874-2.
26. Wahl RL, Jacene H, Kasamon Y, Lodge MA. From RECIST to PERCIST: evolving considerations for PET response criteria in solid tumors. *J Nucl Med Off Pub Soc Nuc Med*. 2009;50 Suppl 1:122S–50. doi:10.2967/jnumed.108.057307.
27. Boellaard R, Delgado-Bolton R, Oyen WJ, Giammarile F, Tatsch K, Eschner W, et al. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. *Eur J Nucl Med Mol Imaging*. 2015;42(2):328–54. doi:10.1007/s00259-014-2961-x.
28. Deasy JO, Blanco AI, Clark VH. CERR: A computational environment for radiotherapy research. *Med Phys*. 2003;30(5):979. doi:10.1118/1.1568978.
29. Parmar C, Rios Velazquez E, Leijenaar R, Jermoumi M, Carvalho S, Mak RH, et al. Robust radiomics feature quantification using semiautomatic volumetric segmentation. *PLoS One*. 2014;9(7), e102107. doi:10.1371/journal.pone.0102107.
30. Skougaard K, Nielsen D, Jensen BV, Hendel HW. Comparison of EORTC criteria and PERCIST for PET/CT response evaluation of patients with metastatic colorectal cancer treated with irinotecan and cetuximab. *J Nucl Med Off Pub Soc Nuc Med*. 2013;54(7):1026–31. doi:10.2967/jnumed.112.111757.
31. Young H, Baum R, Cremerius U, Herholz K, Hoekstra O, Lammertsma AA, et al. Measurement of clinical and subclinical tumour response using [18F]-fluorodeoxyglucose and positron emission tomography: review and 1999 EORTC recommendations. European organization for research and treatment of cancer (EORTC) PET study group. *Eur J Cancer*. 1999;35(13):1773–82.
32. Peeters SG, Zegers CM, Lieuwes NG, van Elmpt W, Eriksson J, van Dongen GA, et al. A comparative study of the hypoxia PET tracers [(1)(8)F]HX4, [(1)(8)F]FAZA, and [(1)(8)F]FMISO in a preclinical tumor model. *Int J Radiat Oncol Biol Phys*. 2015;91(2):351–9. doi:10.1016/j.ijrobp.2014.09.045.
33. Zegers CM, van Elmpt W, Reymen B, Even AJ, Troost EG, Ollers MC, et al. In vivo quantification of hypoxic and metabolic status of NSCLC tumors using [18F]HX4 and [18F]FDG-PET/CT imaging. *Clin Cancer Res Off J Am Assoc Cancer Res*. 2014;20(24):6389–97. doi:10.1158/1078-0432.CCR-14-1524.
34. van der Veldt AA, Smit EF, Lammertsma AA. Positron emission tomography as a method for measuring drug delivery to tumors in vivo: the example of [(11)C]docetaxel. *Frontiers Oncol*. 2013;3:208. doi:10.3389/fonc.2013.00208.
35. van Elmpt W, Ollers M, Dingemans AM, Lambin P, De Ruyscher D. Response assessment using 18F-FDG PET early in the course of radiotherapy correlates with survival in advanced-stage non-small cell lung cancer. *J Nucl Med Off Pub Soc Nuc Med*. 2012;53(10):1514–20. doi:10.2967/jnumed.111.102566.
36. Organization WH. WHO handbook for reporting results of cancer treatment. 1979.
37. Rosner GL, Stadler W, Ratain MJ. Randomized discontinuation design: application to cytostatic antineoplastic agents. *J Clin Oncol Off J Am Soc Clin Oncol*. 2002;20(22):4478–84.
38. Michaelis LC, Ratain MJ. Measuring response in a post-RECIST world: from black and white to shades of grey. *Nat Rev Cancer*. 2006;6(5):409–14. doi:10.1038/nrc1883.
39. Desar IM, van Herpen CM, van Laarhoven HW, Barentsz JO, Oyen WJ, van der Graaf WT. Beyond RECIST: molecular and functional imaging techniques for evaluation of response to targeted therapy. *Cancer Treat Rev*. 2009;35(4):309–21. doi:10.1016/j.ctrv.2008.12.001.
40. Dingemans AM, de Langen AJ, van den Boogaart V, Marcus JT, Backes WH, Scholtens HT, et al. First-line erlotinib and bevacizumab in patients with locally advanced and/or metastatic non-small-cell lung cancer: a phase II study including molecular imaging. *Ann Oncol Off J Eur Soc Med Oncol/ESMO*. 2011;22(3):559–66. doi:10.1093/annonc/mdq391.
41. Hoekstra CJ, Stroobants SG, Smit EF, Vansteenkiste J, van Tinteren H, Postmus PE, et al. Prognostic relevance of response evaluation using [18F]-2-fluoro-2-deoxy-D-glucose positron emission tomography in patients with locally advanced non-small-cell lung cancer. *J Clin Oncol Off J Am Soc Clin Oncol*. 2005;23(33):8362–70. doi:10.1200/JCO.2005.01.1189.