



Primary tumour standardised uptake value is prognostic in nonsmall cell lung cancer: a multivariate pooled analysis of individual data

Marianne Paesmans^{1,18}, Camilo Garcia^{2,18}, Ching-Yee Oliver Wong³, Edward F. Patz Jr⁴, Ritsuko Komaki⁵, Susanne Eschmann⁶, Ramaswamy Govindan⁷, Johan Vansteenkiste⁸, Anne-Pascale Meert⁹, Wouter K. de Jong¹⁰, Nasser Khaled Altorki¹¹, Kotaro Higashi¹², Angela Van Baardwijk¹³, Gerben R. Borst¹⁴, Lieveke Ameye¹, Jean-Jacques Lafitte¹⁵, Thierry Berghmans⁹, Patrick Flamen², Ramon Rami-Porta¹⁷ and Jean-Paul Sculier⁹

Affiliations: ¹Data Centre, Institut Jules Bordet, Université Libre de Bruxelles, Brussels, Belgium. ²Nuclear Medicine, Institut Jules Bordet, Université Libre de Bruxelles, Brussels, Belgium. ³Diagnostic Radiology and Molecular Imaging, Oakland University William Beaumont School of Medicine, William Beaumont Hospital, Royal Oak, MI, USA. ⁴Radiology, Duke University Medical Center, Durham, NC, USA. ⁵The University of Texas MD Anderson Cancer Center, Houston, TX, USA. ⁶Nuclear Medicine, University of Tübingen, Tübingen, Germany. ⁷Division of Medical Oncology, Washington University School of Medicine, St. Louis, MI, USA. ⁸Pneumology, Respiratory Oncology Unit, University Hospitals KU Leuven, Leuven, Belgium. ⁹Intensive Care and Thoracic Oncology, Institut Jules Bordet, Université Libre de Bruxelles, Brussels, Belgium. ¹⁰Pneumology, Ziekenhuis Gelderse Vallei, Ede, The Netherlands. ¹¹Thoracic Surgery, Weill Cornell Medical College, New York, NY, USA. ¹²Radiology, Asanogawa General Hospital, Ishikawa, Japan. ¹³Radiation Oncology (MAASTRO), GROW – School for Oncology and Developmental Biology, Maastricht University Medical Centre, Maastricht, The Netherlands. ¹⁴Radiation Oncology, The Netherlands Cancer Institute, Amsterdam, The Netherlands. ¹⁵Pulmonary and Thoracic Oncology, CHRU de Lille, Lille, France. ¹⁶Nuclear Medicine, CHRU de Lille, Lille, France. ¹⁷Thoracic Surgery, Hospital Universitari Mutua Terrassa and CIBERES Lung Cancer Group, Terrassa, Spain. ¹⁸Both authors contributed equally.

Correspondence: Marianne Paesmans, Data Centre – Institut Jules Bordet, Bd de Waterloo 121, B 1000 Brussels, Belgium. E-mail: marianne.paesmans@bordet.be

ABSTRACT ¹⁸F-fluoro-2-deoxy-D-glucose positron emission tomography (PET) complements conventional imaging for diagnosing and staging lung cancer. Two literature-based meta-analyses suggest that maximum standardised uptake value (SUV_{max}) on PET has univariate prognostic value in nonsmall cell lung cancer (NSCLC). We analysed individual data pooled from 12 studies to assess the independent prognostic value of binary SUV_{max} for overall survival.

After searching the published literature and identifying unpublished data, study coordinators were contacted and requested to provide data on individual patients. Cox regression models stratified for study were used.

Data were collected for 1526 patients (median age 64 years, 60% male, 34% squamous cell carcinoma, 47% adenocarcinoma, 58% stage I–II). The combined univariate hazard ratio for SUV_{max} was 1.43 (95% CI 1.22–1.66) and nearly identical if the SUV threshold was calculated stratifying for histology. Multivariate analysis of patients with stage I–III disease identified age, stage, tumour size and receipt of surgery as independent prognostic factors; adding SUV (HR 1.58, 95% CI 1.27–1.96) improved the model significantly. The only detected interaction was between SUV and stage IV disease.

SUV seems to have independent prognostic value in stage I–III NSCLC, for squamous cell carcinoma and for adenocarcinoma.



@ERSpublications

SUV has independent prognostic value in stage I–III NSCLC, both for squamous cell carcinoma and for adenocarcinoma <http://ow.ly/Q2FKI>

Introduction

Worldwide, lung cancer is the leading cause of cancer-related death in men and the second in women [1]. Despite the availability of new treatments, survival of patients has remained relatively poor [2]. Few validated and accurate prognostic factors are currently used in clinical practice for managing or predicting outcomes for individual patients, although many clinical and histopathological characteristics, laboratory markers, molecular biological markers and gene signatures have been tested for their potential prognostic value. Despite abundant literature on this topic, until now only two characteristics are definitely established as independent prognostic factors: performance status (especially in advanced disease) and disease stage [3], as confirmed by the International Association for the Study of Lung Cancer (IASLC) staging project. In this context, data on >100 000 patients were collected and used to build the 7th edition of the TNM (tumour, node, metastasis) classification for lung cancer [4]. Prognostic factors were also studied, but the findings were limited by the small number of covariates present in the existing worldwide databases. This further indirectly illustrates that very few factors are universally accepted as being important for predicting outcome in lung cancer.

TNM staging assessment was traditionally based on surgical findings and conventional imaging. However, for more than two decades, positron emission tomography (PET) with the glucose analogue 2-¹⁸F-fluoro-2-deoxy-D-glucose (FDG) has been used extensively for disease staging in lung cancer. Indeed, several meta-analyses have shown that the use of FDG-PET improves the accuracy of staging [5, 6].

Our group has performed two systematic reviews and meta-analyses of the literature (the second an update of the first) analysing the possible prognostic value of FDG-PET for survival [7, 8]. The results, published in January 2008 and May 2010, suggested that having a high maximum standardised uptake value (SUV_{max}) of the primary tumour might be a poor prognostic factor. Altogether, 21 studies were analysed [9–29].

In these meta-analyses, our comparison of patients with low and high SUV produced an overall combined hazard ratio of 2.08 for high SUV, significantly different from 1 (95% CI 1.69–2.56). No interaction was detected between older and newer studies ($p=0.60$) or between studies in which patients with metastatic disease were excluded or included ($p=0.46$).

Considering the limitations of literature-based meta-analyses [30], in particular the absence of multivariate analysis, the previous work was extended by performing a pooled analysis of individual patient data, a project endorsed by the IASLC staging project. The primary objective was to assess the prognostic value of primary tumour FDG-PET SUV_{max} (corrected for body weight) for overall survival, with and without adjustment for established prognostic factors, especially the TNM classification. As a secondary objective, subgroup analyses were performed and the possible interactions between SUV and patient or tumour characteristics were investigated.

Methods

All corresponding authors of the studies previously included in the systematic reviews were contacted and provided with a protocol and a plan for further analysing individual data. Those who joined the project were asked to provide such data to the coordinating institution (Institut Jules Bordet, Brussels, Belgium) with pre-specified variables, to update their survival data and to complete a questionnaire about the method used to measure FDG uptake. One unpublished series was also included.

Data from the different series were then reviewed, pooled and analysed at the coordinating institution. A descriptive analysis of the associations between SUV and other possible prognostic factors was performed. Cox regression models were used to obtain, for each individual study, an estimate of the hazard ratio for the effect of SUV on overall survival. SUV was represented for these comparisons as a binary variable (low or high level) with the median in each of the individual series chosen as threshold. The overall impact of SUV was assessed using a stratified Cox model. When possible, overall survival was calculated from the date of PET until the date of death (from any cause) or date of last follow-up. Otherwise, overall survival was used as provided in the individual datasets.

Subgroup analyses were performed using the same method and interaction tests were undertaken. Finally, multivariable Cox regression models stratified by study were used to assess the independent prognostic value of SUV and to assess the benefit of adding SUV to a model based on already-known prognostic variables. For this purpose, a multivariable analysis stratified per study was carried out using the covariates available except SUV. A backward selection method was used. Afterwards, entry of SUV in the model was

This article has supplementary material available from erj.ersjournals.com

Received: Jan 20 2015 | Accepted after revision: July 05 2015 | First published online: Sept 24 2015

Conflict of interest: Disclosures can be found alongside the online version of this article at erj.ersjournals.com

forced to evaluate model improvement. Only patients with complete data were included in the analyses (no missing values). Hazard ratio estimates are reported with 95% confidence intervals, and all p-values are two-tailed. SAS software (version 9.4; SAS Institute Inc., Cary, NC, USA) was used.

The method of measuring SUV and the type of FDG-PET scanner are clearly important, and the reproducibility of measurements between centres depends on the use of standardised protocols. Nevertheless, a large measurement variability is possible and already described elsewhere (see appendixes 1 and 2 of [7]). Briefly, different factors and errors can lead to differences in SUV measurements: clinical/physiological factors (such as glycaemia, time between injection and images, fasting, motion and patient comfort), sources of error (cross-calibration of clocks for correct decay time, accurate cross-calibration of dose calibrator and cameras, exact injected activity and paravenous injection) and physical effects (acquisition settings such as time per bed position, percentage of overlap between bed positions, two- or three-dimensional acquisition mode, type of attenuation correction, image reconstruction methods and image resolution). For these reasons, a specific questionnaire (online supplementary material) was sent to the authors to collect clinical and technical information about the conditions under which both the FDG-PET measurements and the SUV measurements were obtained. The completed questionnaires were intended to be used to construct a covariate, representing subgroups of studies that might be considered homogeneous in the method of acquiring SUV values.

Results

25 corresponding authors were contacted. Five did not reply [14, 21, 22, 26, 31]; another nine [9, 10, 12, 15, 17, 24, 28, 32, 33] did not provide approval for participation or did not send the data after preliminary agreement or did not have the data available at the time we asked them. For the remaining 11 studies [11, 13, 16, 18–20, 23, 25, 27, 29, 34], the individual data were available and received. Additionally, data on one unpublished series of patients was collected.

In terms of numbers of patients, 3578 potentially eligible patients according to the publications (including the unpublished series) were identified and data related to 1563 of these patients (44%) were received. Baseline characteristics of the studies as well as reasons for excluding some patients from the analysis are presented in table 1.

TABLE 1 Baseline characteristics of the studies included in the pooled analysis

First author [ref.]	Year of publication	SUV type	Subjects		Comments	Original threshold	SUV	SUV SQ	SUV AD
			n ₁	n ₂					
AGARWAL [34]	2010	Max.	363	363	Not included in previous meta-analysis	Median	5.9	9.5	4.5
BORST [16]	2005	Max.	51	40	Patients without histological confirmation were excluded	Median	14.3	13.6	13.8
DE JONG [23]	2007	Max.	66	66	Primary author excluded patients with incomplete data	Median	6.0	6.5	6.0
ESCHMANN [20]	2006	Mean	159	136		Best cut-off	9.5	10.3	7.5
GOODGAME [27]	2008	Max.	136	136	Primary author excluded patients	Median	5.5	NR	NR
HIGASHI [13]	2002	Mean	57	48		Best cut-off	2.3		3.0
HOANG [29]	2008	Max./mean	214	221	Some patients were excluded in the original publication	Median	11.1	13.9	9.2
MEERT (unpublished data)		Max.	120	120	Unpublished series [#]	Median	11.7	18.7	9.9
PORT [18]	2005	Max.	64	64		NR	4.04	6.0	2.5
SASAKI [19]	2005	Max.	162	162		Best cut-off	8.2	8.9	6.4
VAN BAARDWIJK [25]	2007	Max.	46	46		Best cut-off	12.8	14.2	12.3
VAN STEENKISTE [11]	1999	Max.	125	124	One patient with SCLC was excluded	Best cut-off	10.3	11.9	7.6

Data are presented as median, unless otherwise stated. SUV: standardised uptake value; n₁: number of patients in the original publication; n₂: number of patients included in the pooled analysis; SQ: squamous cell carcinoma; AD: adenocarcinoma; max.: maximum; NR: not reported; SCLC: small cell lung cancer. [#]: not included in previous meta-analysis.

Patient characteristics, including covariates present in most of the datasets for the 1526 patients included in the analysis are shown in table 2. Plans had been made to adjust the analysis for other known prognostic factors (e.g. biological factors), but this was not possible because of lack of data. Pre-specified analysis of progression-free survival was not possible as the available data allowed calculation of progression-free survival for only ~65% of the patients.

SUV was provided as a continuous covariate in 10 out of the 12 series. In one series, SUV was provided as a binary covariate [27], and in the other, SUV was provided in a mixed way [19]. However, in those two series, the individual hazard ratio could be estimated using the median as the SUV threshold. 10 out of the 12 series reported SUV_{max}, and two did not [13, 20]. In one study [29], both SUV_{max} and SUV_{mean} were available.

Using ANOVAs adjusted for study, an association between SUV (as a continuous covariate) and sex, with higher metabolic activity (and hence SUV) in male patients ($p=0.04$) was identified. In contrast, metabolic activity was lower in adenocarcinoma ($p<0.001$) (as illustrated in table 1 and fig. 1) and in stage I or stage II disease ($p<0.001$). Finally, a significant positive correlation was noted between SUV and tumour size (both variables continuous, $p<0.001$), but not with age.

The impact of SUV on overall survival is shown in table 3, together with the rates of events and the hazard ratios we obtained in our literature-based meta-analysis. The main differences between the literature-based hazard ratio estimates and the presently obtained hazard ratio estimates resulted from the use of SUV_{median} as the threshold or to updates of survival data or to slight changes in the populations of patients analysed.

The combined hazard ratio was 1.43 (95% CI 1.22–1.66): patients with a high SUV had significantly worse survival than patients with a low SUV. If the two studies in which SUV_{mean} was used were excluded, the combined hazard ratio became 1.46 (95% CI 1.23–1.73), with the same conclusion. Survival curves are depicted in figure 2.

When SUV was dichotomised according to the median values with stratification by histology (adenocarcinoma *versus* squamous cell *versus* other histological subtype), 159 (10%) patients changed SUV categories (88 from a low value to a high value and 71 in the other direction). However, the combined hazard ratio did not change appreciably (HR 1.43, 95% CI 1.23–1.67).

Results of subgroup analyses are shown in table 4. The only interaction identified was with stage, suggesting that SUV has little or no prognostic value among patients with stage IV disease.

The first fitted multivariable model included age, stage, tumour size and receipt of surgery as explanatory variables (981 patients and 337 events). Because sex was not in the model, refitting was performed, adding patients with unknown sex category (1117 patients and 374 events). Older age (HR 0.51, 95% CI 0.40–0.66; $p<0.0001$), stage I (HR 0.42, 95% CI 0.29–0.60; $p<0.0001$) or stage II (HR 0.60, 95% CI 0.40–0.91; $p<0.0001$) were favourable prognostic factors; larger tumour size (HR 1.09, 95% CI 1.04–1.14; $p=0.0006$) and lack of surgery (HR 2.98, 95% CI 2.04–4.35; $p<0.0001$) were poor prognostic factors. Notably, because

TABLE 2 Patient characteristics in the individual series

First author [ref.]	Age [#] years	Male [¶]	Tumour histology			Disease stage				Tumour size ⁺ cm median	Surgery
			SQ	AD	Other	I	II	III	IV		
AGARWAL [34]	49% aged <70 years	48	25	63	13	92	8			2.2	100
BORST [16]	71 [31–87]	65	43	33	25	38	20	43		4.8	0
DE JONG [23]	62 [38–79]	76	53	35	13	59	27	14		3.1	100
ESCHMANN [20]	58 [30–78]	76	47	29	24			100		5.0	60
GOODGAME [27]	68 [33–83]	NR	35	48	18	100				2.5	100
HIGASHI [13]	65 [46–85]	44		73	27	83	2	15		2.5	100
HOANG [29]	64 [37–89]	61	29	36	36			55	45	NR	22
MEERT (unpublished data)	62 [29–86]	66	25	53	23	12	4	28	57	NR	10
PORT [18]	68 [46–81]	41	8	88	5	81	8	11		1.5	94
SASAKI [19]	62 [34–88]	48	43	44	14	40	16	44		3.6	57
VAN BAARDWIJK [25]	67 [43–85]	67	48	39	13	54	17	28		3.5	100
VANSTEENKISTE [11]	64 [33–83]	89	55	23	23	37	15	48		3.9	73

Data are presented as median (range) or %, unless otherwise stated. SQ: squamous cell carcinoma; AD: adenocarcinoma; NR: not reported. [#]: age was provided as a continuous variable in 11 studies and as a binary variable in one study; [¶]: patient sex was provided in 11 out of the 12 series; ⁺: tumour size was provided in 10 out of the 12 series.

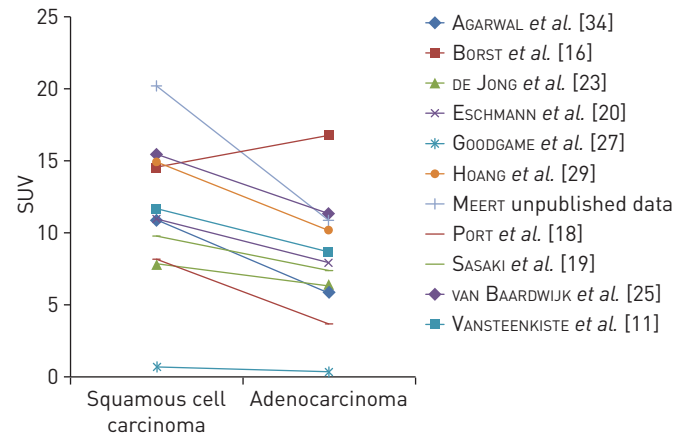


FIGURE 1 Marginal means for standardised uptake value (SUV) for the subgroups of patients with squamous cell carcinoma or adenocarcinoma by study.

tumour size was not available for any patient with stage IV disease, this model did not include any stage IV patient. Addition of SUV to this previous model significantly improved it (table 5). The hazard ratio for high *versus* low SUV was 1.58 (95% CI 1.27–1.96; $p < 0.0001$) and increased slightly if the threshold used to categorise it depended on histology (HR 1.61, 95% CI 1.30–2.00). No interaction between SUV and other covariates included in the model was found to be significant.

When tumour size was not considered, stage IV patients could be added to the analysis, again without statistical significance for sex. Without SUV, the best model (1526 patients and 669 events) included age, stage and lack of surgery, with the following hazard ratios. Age ≥ 70 years: 0.66 (95% CI 0.55–0.78; $p < 0.0001$); stage I: 0.20 (95% CI 0.13–0.29; $p < 0.0001$); stage II: 0.31 (95% CI 0.20–0.48; $p < 0.0001$); stage III: 0.48 (95% CI 0.37–0.62; $p < 0.0001$); and lack of surgery: 2.46 (95% CI 1.94–3.13; $p < 0.0001$). The addition of SUV to this model improved its predictive ability, as shown in table 5 (similar results for the two different categorisations of SUV). In this latter case, the interaction between stage and SUV was significant with no detectable effect of SUV in patients with stage IV.

The data collected on the methods used to obtain SUV are shown in table 6. This table shows that the various studies did not have identical clinical or physiological patient preparations, probably did not have the same physical effects and had some unassessable sources of error because several studies were retrospective. This substantial heterogeneity did not allow us to construct groups of studies as planned, and therefore prevented us from studying SUV as a continuous variable.

Discussion

We performed a pooled analysis based on individual data from 1526 patients from 12 different series; in other words, 44% of the patients identified as eligible were included in the analysis. We confirmed the results of our previous work: patients with high SUV uptake of the primary tumour have shorter survival than patients with lower metabolic rate, and we have now identified high SUV as an independent

TABLE 3 Effect of standardised uptake value (SUV) on overall survival: univariate analysis

First author [ref.]	Original threshold	Literature-based	Individual patient data (original threshold)	Individual patient data (SUV _{median} as threshold)
AGARWAL [34]	Median	NA		1.87 (1.13–3.09)
BORST [16]	Median	3.15 (1.59–6.22)		3.84 (1.66–8.90)
DE JONG [23]	Median	2.93 (1.21–7.09)		3.27 (1.35–7.89)
ESCHMANN [20]	Best cut-off	1.71 (1.00–2.93)	1.59 (1.02–2.47)	1.06 (0.71–1.59)
GOODGAME [27]	Median	1.89 (1.20–2.99)		1.88 (0.97–3.64)
HIGASHI [13]	Best cut-off	6.20 (1.34–28.75)		11.79 (1.49–93.24)
HOANG [29]	Median	1.29 (0.94–1.76)		1.28 (0.94–1.75)
MEERT (unpublished data)	Median	NA		0.95 (0.62–1.45)
PORT [18]	Literature	2.36 (0.24–22.88)	2.98 (0.33–26.78)	0.56 (0.09–3.41)
SASAKI [19]	Best cut-off	7.66 (1.41–41.50)	5.74 (1.36–24.23)	1.39 (0.66–2.94)
VAN BAARDWIJK [25]	Best cut-off	3.40 (1.40–8.26)	2.20 (0.95–5.10)	1.54 (0.77–3.09)
VANSTEENKISTE [11]	Best cut-off	2.72 (1.50–4.94)	2.01 (1.21–3.33)	1.42 (0.95–2.11)

Data are presented as HR (95% CI), unless otherwise stated. NA: not applicable (study was not included in the literature-based meta-analyses).

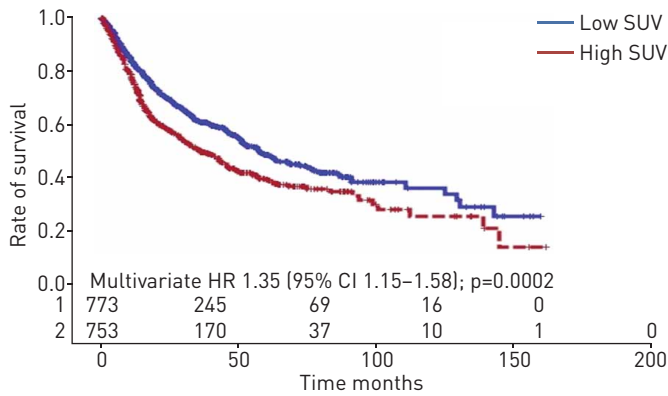


FIGURE 2 Overall survival according to standardised uptake value (SUV), dichotomised at the median in each dataset. 1: low SUV; 2: high SUV.

prognostic factor, a conclusion that could not be reached from our previous literature-based meta-analysis. However, as expected, the point estimates were lower than those previously obtained; indeed, we used the SUV_{median} values in each series as threshold to dichotomise SUV measurements instead of the so-called “best cut-off”, which is known to increase the chance of false-positive results [35]; adjustment for other prognostic covariates may also explain the lower point estimates. As expected, tumour metabolic activity was associated with histology. To correct for this factor when assessing the prognostic value of SUV, we calculated separate thresholds stratifying by histology, but we did not identify any modification of the magnitude of SUV’s effect on survival by doing so. Similarly, we did not find an interaction between SUV and histology. This is also a new finding obtained thanks to the use of individual patient data.

There are several limitations to this study. The number of covariates was restricted compared with our protocol; as previously discussed [3, 4], this illustrates the need to standardise the study of prognostic factors in lung cancer. One of the cited advantages of an individual patient data meta-analysis [30] is to provide more mature data, but in the present analysis we observed that it is difficult, at least outside the context of clinical trials, to perform searches to update survival status. Finally, we did not have a sufficient number of datasets that included information on disease progression, and thus we could not calculate progression-free survival, which would have been a further advantage over our previous systematic reviews.

TABLE 4 Subgroup analyses

Subgroup	Patients n	HR (95% CI)	p-value for interaction
Sex			
Female	557	1.58 (1.18–2.10)	0.67
Male	833	1.39 (1.14–1.69)	
Age			
<70 years	1000	1.30 (1.07–1.57)	0.09
≥70 years	526	1.70 (1.28–2.24)	
Histology			
SQ	512	1.47 (1.12–1.94)	0.62
AD	718	1.52 (1.19–1.95)	
LC	296	1.30 (0.94–1.79)	
SQ [#]	512	1.42 (1.10–1.84)	
AD [#]	718	1.53 (1.20–1.95)	
LC [#]	296	1.31 (0.95–1.80)	
Disease stage			
I	766	1.60 (1.18–2.17)	0.05
II	117	2.59 (1.25–5.39)	
III	475	1.31 (1.04–1.64)	
IV	168	0.98 (0.70–1.38)	
Surgery			
No	481	1.25 (1.00–1.55)	0.15
Yes	1045	1.55 (1.25–1.94)	

SQ: squamous cell carcinoma; AD: adenocarcinoma; LC: large cell carcinoma. #: threshold for dichotomising SUV dependent on histology (per study).

TABLE 5 Multivariable models

Covariate	HR (95% CI)	p-value
Excluding patients with stage IV cancer		
Age (ref.: ≥ 70 years)	0.49 (0.38–0.63)	<0.0001
Stage I (ref.: stage III)	0.41 (0.29–0.59)	<0.0001
Stage II (ref.: stage III)	0.57 (0.37–0.86)	
Tumour size	1.06 (1.01–1.11)	0.02
No surgical treatment	3.14 (2.14–4.61)	<0.0001
SUV (ref.: low)	1.58 (1.27–1.96)	<0.0001
Including patients with stage IV cancer		
Age (ref.: ≥ 70 years)	0.65 (0.55–0.78)	<0.0001
Stage I (ref.: stage IV)	0.21 (0.14–0.31)	<0.0001
Stage II (ref.: stage IV)	0.31 (0.20–0.48)	
Stage III (ref.: stage IV)	0.48 (0.37–0.62)	
No surgical treatment	2.45 (1.93–3.11)	<0.0001
SUV (ref.: low)	1.35 (1.15–1.58)	0.0002

Ref.: reference; SUV: standardised uptake value.

Our literature-based meta-analyses [7, 8] prevented us from addressing the issue of heterogeneity of studies with regards to the method of assessing SUV and of the subsequent large variability in measurements. This is why our analyses of correlations between SUV as a continuous variable and other baseline characteristics should be interpreted carefully, although they were adjusted for the study. Some scientific associations, such as the European Association of Nuclear Medicine have not only published guidelines for semi-quantitative FDG-PET evaluation in oncology [36], but have also built a system for PET/computed tomography centre accreditation [37], taking into account staff training, clinical conditions of the examination, calibration of the machines, methods of reconstruction and attenuation algorithms, among others. However, accreditation does not fully solve the issue of reproducibility of results [38]. By using a questionnaire asking specifically about the methods used in each study, we were unable to identify homogeneous groups, and we can only conclude that higher values of SUV imply higher hazards. Some authors have suggested that patients do not fall into two discrete groups, but rather show a continuous increase in the hazard as SUV increases [23, 39]. However, our data did not allow us to propose one or several thresholds. In this field of metabolic imaging perhaps more than in any other, prospective standardised multicentre studies with an *a priori* estimated sample size allowing sufficient statistical power are lacking but greatly needed.

Publication bias is another important issue, and clearly our study is likely to have suffered from such a bias: we missed 56% of the “eligible” patients, and many other series may well exist without being published. This is why our study has to be interpreted as a pooled data analysis and is not a meta-analysis of the available evidence. With the limited number of patients, there is no power issue, as a significant independent prognostic value for high SUV was identified, but there is a problem of publication bias, intrinsic to any meta-analysis of prognostic factors, as no registry of such studies exists. This is the price to pay when conducting individual patient data meta-analyses, as obtaining approval to share data from all database owners is very difficult; we partially failed in this exercise despite the fact that we shared with all the authors contacted a protocol for the present study and a pre-planned analysis. However, when returning to our literature-based meta-analysis and calculating combined hazard ratios for studies with individual patient data available or not, similar point estimates were found: 2.15 (95% CI 1.57–2.93) and 2.03 (95% CI 1.54–2.69) with a p-value of 0.79 for heterogeneity test. This suggests that our subgroup of studies might be well representative of the overall evidence and that publication bias is not important. As discussing data exchange with the authors of published series takes a lot of time, we did not try to obtain more recently published data. However, should high SUV have truly no effect at all on survival, a very large study (or many medium-sized studies) would be required to transform our results into non-statistically significant ones.

Nevertheless, for all the above reasons, a prospective study aiming to detect a hazard ratio similar to the point estimate we obtained would be necessary to confirm our results.

Despite the mixture of SUV measurements, only two series provided SUV_{mean} , compared to 10 series that measured SUV_{max} . The exclusion of these two series did not significantly affect the results nor modify the conclusion that patients with a high SUV have significantly worse survival than patients with a low SUV.

TABLE 6 Methods of standardised uptake value (SUV) assessment

First author [ref.]	Type of camera	Trademark and model	Estimated clinical spatial resolution of PET camera (FWHM)	Fasting	Injected activity	Acquisition time per bed position (emission/transmission)	Quantification value used	SUV normalisation	Incubation time between injection and image acquisition	SUV obtained from clinical report versus specific review	Glycaemia levels accepted
AGARWAL [34]	PET and PET/CT	GE Discovery DST and GE Advance	5 mm (NA for PET alone)	4 h	370 MBq average	3 min emission (NA for PET alone)	SUV _{max}	Body weight	60 min	Report (a minority review)	NA
BORST [16]	PET	Siemens/CTI	8 mm	>6 h	350–550 MBq	10–15 min emission/transmission	SUV _{max}	Body weight and glucose corrected	60 min	Review	No limit
DE JONG [23]	PET	Siemens ECAT (two centres)	5 mm	>6 h	220–370 MBq	5 min emission/3 min transmission	SUV _{max}	Body weight	60–90 min	Both	NA
ESCHMANN [20]	PET	GE Advance	5 mm	>12 h	350–450 MBq	NA	SUV _{max} and SUV _{mean}	Body weight	60 min	Both	150 mg·dL ⁻¹
GOODGAME [27]	PET and PET/CT	Siemens ECAT Exact, Philips/ADAC CPET Plus and Siemens Biograph LSO Duo Hybrid	NA	NA	NA	NA	SUV _{max}	NA	NA	NA	NA
HIGASHI [13]	PET	Shimazu Headtome IV	5 mm	>4 h	185 MBq	5 min emission/5 min transmission	SUV _{mean}	Body weight	40 min	Clinical report	120 mg·dL ⁻¹
HOANG [29]	PET	GE 4096 plus and GE Advance Unit	5 mm	4 h	370–555 MBq	NA	SUV _{max}	Body weight	60±15 min	Review	<200 mg·dL ⁻¹
MEERT (unpublished data)	PET/CT	GE Discovery LS	8 mm	>6 h	3.7–7.4 MBq·kg ⁻¹	4 min emission	SUV _{max}	Body weight	60–75 min	Review	150 ml·dL ⁻¹ l
PORT [18]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
SASAKI [19]	PET	CTI Siemens HR	8 mm	>6 h	370–555 MBq	5 min emission/3 min transmission	SUV _{max}	Body weight	60 min	Review	150 mg·dL ⁻¹
VAN BAARDWIJK [25]	PET	Siemens ECAT EXACT 922 and CTI Siemens 931/08/12	7 mm	>6 h	3.5 MBq·kg ⁻¹	8 min emission/4 min transmission	SUV _{max}	Body weight	60 min (range 45–120 min)	Both	~150 mg·dL ⁻¹
VANSTEENKISTE [11]	PET	CTI Siemens 931/08/12	8 mm	>6 h	6.5 MBq·kg ⁻¹	NA	SUV _{max}	Body weight	50–70 min	Review	NA

PET: positron emission tomography; FWHM: full width at half maximum; CT: computed tomography; NA: no data available.

Interestingly, we confirmed that lung adenocarcinomas showed the lowest SUV values among tumour histologies, as has been reported by several authors [32, 40–42]. As far as the analysis could be taken without using SUV as a continuous covariate, this correlation did not modify the prognostic value of SUV.

The well-known partial volume effect [43] could at least partially explain the finding that patients with stage I–II lung cancer showed lower metabolic uptake. Nevertheless, SUV consistently showed prognostic value in each stage, from stage I to III, but our results suggest that SUV is not prognostic among patients with stage IV disease. It is possible that once a lung tumour shows metastatic spread, the anatomic extent of the tumour is more important in predicting prognosis than the metabolic activity of the primary tumour only. Although this supposition should be confirmed with a larger number of patients with stage IV disease, this might be the reason why the introduction of other measures of metabolic activity such as tumour glycolysis and metabolic volume as well as heterogeneity and texture analysis are very promising as prognostic factors [44–49]. However, these “new” derived metabolic parameters seem to be more interesting in large tumours, due to the inability of PET to characterise tracer distribution within small tumours because of its limited spatial resolution. However, multiple texture parameters have been described and tested to describe the tumour heterogeneity but no consensus has been reached to conclude which is the best and the most robust. In addition, these parameters are also very interesting in NSCLC stage IV patients, where we failed to identify any independent prognostic value for SUV_{max}. In our opinion, multicentric large prospective trials testing these new derived parameters are needed to conclude in which cases/stages their use is the most pertinent.

Conclusions

Although our study suffered from selection bias and lack of standardised SUV assessment, our findings suggest that SUV at the time of diagnosis, measured on the primary tumour, is an independent prognostic marker for patients with stage I–III NSCLC. There was no detectable interaction between histology and SUV, despite the confirmed observation that SUV uptake is higher in squamous cell cancer compared to adenocarcinoma. The utility of SUV in predicting survival in stage IV patients requires further study.

Acknowledgements

The authors thank Claude Hossein-Foucher (Nuclear Medicine, CHRU de Lille, Lille, France) and Arnaud Scherpereel (Pulmonary and Thoracic Oncology, CHRU de Lille) for their support of the work and their critical reading of the manuscript.

References

- 1 International Agency for Research on Cancer. Estimated Cancer Incidence, Mortality and Prevalence Worldwide in 2012. Population Fact Sheets. http://globocan.iarc.fr/Pages/fact_sheets_population.aspx Date last accessed: August 5, 2014.
- 2 Youlden DR, Cramb SM, Baade PD. The international epidemiology of lung cancer: geographical distribution and secular trends. *J Thorac Oncol* 2008; 3: 819–831.
- 3 Paesmans M. Facteurs pronostiques du cancer bronchique. [Prognostic factors in lung cancer]. *Rev Mal Respir* 2005; 22: 8S76–8S80.
- 4 Sculier JP, Chansky K, Crowley JJ, et al. The impact of additional prognostic factors on survival and their relationship with the anatomical extent of disease expressed by the 6th edition of the TNM Classification of Malignant Tumors and the proposals for the 7th edition. *J Thorac Oncol* 2008; 3: 457–466.
- 5 Gould MK, Kuschner WG, Rydzak CE, et al. Test performance of positron emission tomography and computed tomography for mediastinal staging in patients with non-small-cell lung cancer: a meta-analysis. *Ann Intern Med* 2003; 139: 879–892.
- 6 de Langen AJ, Raijmakers P, Riphagen I, et al. The size of mediastinal lymph nodes and its relation with metastatic involvement: a meta-analysis. *Eur J Cardiothorac Surg* 2006; 29: 26–29.
- 7 Berghmans T, Dusart M, Paesmans M, et al. Primary tumor standardized uptake value (SUV_{max}) measured on fluorodeoxyglucose positron emission tomography (FDG-PET) is of prognostic value for survival in non-small cell lung cancer (NSCLC): a systematic review and meta-analysis (MA) by the European Lung Cancer Working Party for the IASLC Lung Cancer Staging Project. *J Thorac Oncol* 2008; 3: 6–12.
- 8 Paesmans M, Berghmans T, Dusart M, et al. Primary tumor standardized uptake value measured on fluorodeoxyglucose positron emission tomography is of prognostic value for survival in non-small cell lung cancer: update of a systematic review and meta-analysis by the European Lung Cancer Working Party for the International Association for the Study of Lung Cancer Staging Project. *J Thorac Oncol* 2010; 5: 612–619.
- 9 Ahuja V, Coleman RE, Herndon J, et al. The prognostic significance of fluorodeoxyglucose positron emission tomography imaging for patients with nonsmall cell lung carcinoma. *Cancer* 1998; 83: 918–924.
- 10 Sugawara Y, Quint LE, Iannettoni MD, et al. Does the FDG uptake of primary non-small cell lung cancer predict prognosis? A work in progress. *Clin Positron Imaging* 1999; 2: 111–118.
- 11 Vansteenkiste JF, Stroobants SG, Dupont PJ, et al. Prognostic importance of the standardized uptake value on ¹⁸F-fluoro-2-deoxy-glucose-positron emission tomography scan in non-small-cell lung cancer: an analysis of 125 cases. *J Clin Oncol* 1999; 17: 3201–3206.
- 12 Dhital K, Saunders CA, Seed PT, et al. [¹⁸F]Fluorodeoxyglucose positron emission tomography and its prognostic value in lung cancer. *Eur J Cardiothorac Surg* 2000; 18: 425–428.
- 13 Higashi K, Ueda Y, Arisaka Y, et al. ¹⁸F-FDG uptake as a biologic prognostic factor for recurrence in patients with surgically resected non-small cell lung cancer. *J Nucl Med* 2002; 43: 39–45.

- 14 Jeong HJ, Min JJ, Park JM, *et al.* Determination of the prognostic value of [¹⁸F]fluorodeoxyglucose uptake by using positron emission tomography in patients with non-small cell lung cancer. *Nucl Med Commun* 2002; 23: 865–870.
- 15 Downey RJ, Akhurst T, Gonen M, *et al.* Preoperative F-18 fluorodeoxyglucose-positron emission tomography maximal standardized uptake value predicts survival after lung cancer resection. *J Clin Oncol* 2004; 22: 3255–3260.
- 16 Borst GR, Belderbos JS, Boellaard R, *et al.* Standardised FDG uptake: a prognostic factor for inoperable non-small cell lung cancer. *Eur J Cancer* 2005; 41: 1533–1541.
- 17 Cerfolio RJ, Bryant AS, Ohja B, *et al.* The maximum standardized uptake values on positron emission tomography of a non-small cell lung cancer predict stage, recurrence, and survival. *J Thorac Cardiovasc Surg* 2005; 130: 151–159.
- 18 Port JL, Andrade RS, Levin MA, *et al.* Positron emission tomographic scanning in the diagnosis and staging of non-small cell lung cancer 2 cm in size or less. *J Thorac Cardiovasc Surg* 2005; 130: 1611–1615.
- 19 Sasaki R, Komaki R, Macapinlac H, *et al.* [¹⁸F]fluorodeoxyglucose uptake by positron emission tomography predicts outcome of non-small-cell lung cancer. *J Clin Oncol* 2005; 23: 1136–1143.
- 20 Eschmann SM, Friedel G, Paulsen F, *et al.* Is standardised ¹⁸F-FDG uptake value an outcome predictor in patients with stage III non-small cell lung cancer? *Eur J Nucl Med Mol Imaging* 2006; 33: 263–269.
- 21 Prévost S, Boucher L, Larivée P, *et al.* Bone marrow hypermetabolism on ¹⁸F-FDG PET as a survival prognostic factor in non-small cell lung cancer. *J Nucl Med* 2006; 47: 559–565.
- 22 Raz DJ, Odisho AY, Franc BL, *et al.* Tumor fluoro-2-deoxy-D-glucose avidity on positron emission tomographic scan predicts mortality in patients with early-stage pure and mixed bronchioloalveolar carcinoma. *J Thorac Cardiovasc Surg* 2006; 132: 1189–1195.
- 23 de Jong WK, van der Heijden HF, Pruijm J, *et al.* Prognostic value of different metabolic measurements with fluorine-18 fluorodeoxyglucose positron emission tomography in resectable non-small cell lung cancer: a two-center study. *J Thorac Oncol* 2007; 2: 1007–1012.
- 24 Downey RJ, Akhurst T, Gonen M, *et al.* Fluorine-18 fluorodeoxyglucose positron emission tomographic maximal standardized uptake value predicts survival independent of clinical but not pathologic TNM staging of resected non-small cell lung cancer. *J Thorac Cardiovasc Surg* 2007; 133: 1419–1427.
- 25 van Baardwijk A, Dooms C, van Suylen RJ, *et al.* The maximum uptake of ¹⁸F-deoxyglucose on positron emission tomography scan correlates with survival, hypoxia inducible factor-1 α and GLUT-1 in non-small cell lung cancer. *Eur J Cancer* 2007; 43: 1392–1398.
- 26 Zhang ZJ, Chen JH, Meng L, *et al.* ¹⁸F-FDG uptake as a biologic factor predicting outcome in patients with resected non-small-cell lung cancer. *Chin Med J (Engl)* 2007; 120: 125–131.
- 27 Goodgame B, Pillot GA, Yang Z, *et al.* Prognostic value of preoperative positron emission tomography in resected stage I non-small cell lung cancer. *J Thorac Oncol* 2008; 3: 130–134.
- 28 Hanin FX, Lonnew M, Cornet J, *et al.* Prognostic value of FDG uptake in early stage non-small cell lung cancer. *Eur J Cardiothorac Surg* 2008; 33: 819–823.
- 29 Hoang JK, Hoagland LF, Coleman RE, *et al.* Prognostic value of fluorine-18 fluorodeoxyglucose positron emission tomography imaging in patients with advanced-stage non-small-cell lung carcinoma. *J Clin Oncol* 2008; 26: 1459–1464.
- 30 Piedbois P, Buyse M. Meta-analyses based on abstracted data: a step in the right direction, but only a first step. *J Clin Oncol* 2004; 22: 3839–3841.
- 31 Na II, Cheon GJ, Choe DH, *et al.* Clinical significance of ¹⁸F-FDG uptake by N2 lymph nodes in patients with resected stage IIIA N2 non-small-cell lung cancer: a retrospective study. *Lung Cancer* 2008; 60: 69–74.
- 32 Vesselle H, Schmidt RA, Pugsley JM, *et al.* Lung cancer proliferation correlates with [F-18]fluorodeoxyglucose uptake by positron emission tomography. *Clin Cancer Res* 2000; 6: 3837–3844.
- 33 Kieninger AN, Welsh R, Bendick PJ, *et al.* Positron-emission tomography as a prognostic tool for early-stage lung cancer. *Am J Surg* 2006; 191: 433–436.
- 34 Agarwal M, Brahmanday G, Bajaj SK, *et al.* Revisiting the prognostic value of preoperative ¹⁸F-fluoro-2-deoxyglucose (¹⁸F-FDG) positron emission tomography (PET) in early-stage (I & II) non-small cell lung cancers (NSCLC). *Eur J Nucl Med Mol Imaging* 2010; 37: 691–698.
- 35 Altman DG, Lausen B, Sauerbrei W, *et al.* Dangers of using "optimal" cutpoints in the evaluation of prognostic factors. *J Natl Cancer Inst* 1994; 86: 829–835.
- 36 Boellaard R, O'Doherty MJ, Weber WA, *et al.* FDG PET and PET/CT: EANM procedure guidelines for tumour PET imaging: version 1.0. *Eur J Nucl Med Mol Imaging* 2010; 37: 181–200.
- 37 FDG-PET/CT Accreditation. http://earl.eanm.org/cms/website.php?id=en/projects/fdg_pet_ct_accreditation.htm Date last accessed: August 5, 2014.
- 38 Boellaard R, Oyen WJ, Hoekstra CJ, *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: 2320–2333.
- 39 Detterbeck FC, Vansteenkiste JF, Morris DE, *et al.* Seeking a home for a PET, part 3: emerging applications of positron emission tomography imaging in the management of patients with lung cancer. *Chest* 2004; 126: 1656–1666.
- 40 Higashi K, Ueda Y, Seki H, *et al.* Fluorine-18-FDG PET imaging is negative in bronchioloalveolar lung carcinoma. *J Nucl Med* 1998; 39: 1016–1020.
- 41 Brown RS, Leung JY, Kison PV, *et al.* Glucose transporters and FDG uptake in untreated primary human non-small cell lung cancer. *J Nucl Med* 1999; 40: 556–565.
- 42 Kim BT, Kim Y, Lee KS, *et al.* Localized form of bronchioloalveolar carcinoma: FDG PET findings. *AJR Am J Roentgenol* 1998; 170: 935–939.
- 43 Soret M, Bacharach SL, Buvat I. Partial-volume effect in PET tumor imaging. *J Nucl Med* 2007; 48: 932–945.
- 44 Liao S, Penney BC, Wroblewski K, *et al.* Prognostic value of metabolic tumor burden on ¹⁸F-FDG PET in nonsurgical patients with non-small cell lung cancer. *Eur J Nucl Med Mol Imaging* 2012; 39: 27–38.
- 45 Zaizen Y, Azuma K, Kurata S, *et al.* Prognostic significance of total lesion glycolysis in patients with advanced non-small cell lung cancer receiving chemotherapy. *Eur J Radiol* 2012; 81: 4179–4184.
- 46 Orhac F, Soussan M, Maisonobe JA, *et al.* Tumor texture analysis in ¹⁸F-FDG-PET: relationships between texture parameters, histogram indices, SUVs, metabolic volumes, and total lesion glycolysis. *J Nucl Med* 2014; 55: 414–422.

- 47 Chen HH, Chiu NT, Su WC, *et al.* Prognostic value of whole-body total lesion glycolysis at pretreatment FDG PET/CT in non-small cell lung cancer. *Radiology* 2012; 264: 559–566.
- 48 Tixier F, Hatt M, Valla C, *et al.* Visual versus quantitative assessment of intratumor ¹⁸F-FDG PET uptake heterogeneity: prognostic value in non-small cell lung cancer. *J Nucl Med* 2014; 55: 1235–1241.
- 49 Cook GJ, Yip C, Siddique M, *et al.* Are pretreatment ¹⁸F-FDG PET tumor textural features in non-small cell lung cancer associated with response and survival after chemoradiotherapy? *J Nucl Med* 2013; 54: 19–26.