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

Thierry Berghmans, MD, Michèle Dusart, MD,† Marianne Paesmans, MSc,‡ Claude Hossein-Foucher, MD,§ Irene Buvat, PhD,‖ Catherine Castaigne, MD,¶ Arnaud Scherpereel, MD, PhD,# Céline Mascaux, MD,* ** Michel Moreau, MD,‡ Martine Roelandts, MD,†† Stéphane Alard, MD,‡‡ Anne-Pascale Meert, MD, PhD,* Edward F. Patz, Jr, MD,§§‖‖ Jean-Jacques Lafitte, MD, PhD,# and Jean-Paul Sculier, MD, PhD*‖‖*

Hypothesis: The 2-[^{18}F]-fluoro-2-deoxy-D-glucose positron emission tomography is an imaging tool for assessing clinical tumor, node, metastasis in non-small cell lung cancer (NSCLC). Primary tumor standardized uptake value (SUV) has been studied as a potential prognostic factor for survival. However, the sample sizes are limited leading to conduct a meta-analysis to improve the precision in estimating its effect.

Methods: We performed a systematic literature search. For each publication, we extracted an estimate of the hazard ratio (HR) for comparing patients with a low and a high SUV and we aggregated the individual HRs into a combined HR, using a random-effects model.

Results: We found 13 eligible studies dedicated to NSCLC. Most of them included patients with stages I to III/IV and used a SUV assessment corrected for body weight. Number of patients ranged from 38 to 315 (total: 1474); 11 studies identified a high SUV as a poor prognostic factor for survival although two studies found no significant correlation between SUV and survival. SUV measurement and SUV threshold for defining high SUV were study dependent, eight studies looked for a so-called best cutoff (maximizing the logrank test statistic) without adjusting the *p* value for multiplicity. Overall, the combined HR for the 13 reports was 2.27 (95% confidence interval [CI]: 1.70–3.02); excluding the studies proposing a "best" cutoff, it was 2.08 (95% CI: 1.431–3.04).

Conclusion: Our meta-analysis suggests that the primary tumor SUV measurement has a prognostic value in NSCLC; these results should be confirmed in a meta-analysis on individual patients' data.

Key Words: SUV, Meta-analysis, Non-small cell lung cancer, PET scan.

(*J Thorac Oncol.* 2008;3: 6–12)

With the exception of stage and performance status, no prognostic factors have been definitively established in lung cancer.[1,2] Clinical features including gender, age, weight loss, and serum markers, such as lactate dehydrogenase, neuron-specific enolase, cytokeratin fragment 21-1 levels, or leukocytes or neutrophils counts, have been studied but are not sufficiently accurate for individual patient management.[2,3] More accurate markers would be helpful to stratify patients for therapy and predict outcomes.

Cancer stage is currently the most important prognostic factor for survival, also having implications in the therapeutic strategy. The last, and 6th edition of the *TNM Classification of Malignant Tumors*, mainly based on surgical patients, was published in 1997.[4] This staging system relied on conventional imaging (cases collected from 1975), but during the last decade, positron emission tomography (PET) with the glucose analogue 2-[^{18}F]-fluoro-2-deoxy-D-glucose (^{18}F-

*Department of Intensive Care Unit and Thoracic Oncology; †Department of Radionuclide Imaging; ‡Data Centre, and ††Department of Radiotherapy, Institut Jules Bordet, Université Libre de Bruxelles (ULB), Bruxelles, Belgium; §Department of Radionuclide Imaging and #Department of Pneumology, CHRU de Lille, Lille, France; ‖UMR 678 INSERM-UPMC, CHU Pitié-Salpêtrière, Paris, France; ¶Department of Radionuclide Imaging and ‡‡Department of Radiology, Hôpital Saint-Pierre, Bruxelles, Belgium; **Research Fellow FNRS, §§Departments of Radiology, Pharmacology and Cancer Biology, Duke University Medical Center, Durham, North Carolina; ‖‖IASLC Lung Cancer Staging Committee.

Disclosure: The authors declare no conflict of interest.

Address for correspondence: Thierry Berghmans, MD, Institut Jules Bordet, Rue Héger-Bordet, 1, 1000 Bruxelles, Belgium. E-mail: thierry.berghmans@bordet.be

ISSN: 1556-0864/08/0301-0006

FDG) has emerged as an useful tool for several malignancies. For lung cancer, two meta-analyses[5,6] have established the superiority of ^{18}F-FDG-PET over computed tomography (CT) and other imaging modalities for the diagnosis and staging of lung cancer.

Other retrospective studies have explored the survival prognostic significance of the standardized uptake value (SUV) value, a semiquantitative simplified measurement of tissue deoxyglucose metabolic rate, but most of these reports only include a small number of patients. Based on these considerations and in the context of the International Association for the Study of Lung Cancer (IASLC) Staging Project proposals for the forthcoming, 7th edition of the *TNM Classification for Malignant Tumors*, we performed a systematic review of the literature on ^{18}F-FDG-PET scan and survival and a meta-analysis of the data to determine the prognostic value of primary tumor SUV in patients with lung cancer.

MATERIALS AND METHODS

To be eligible for the systematic review, a study had to fulfill the following criteria: limited to lung cancer (any stage or any histology), ^{18}F-FDG-PET studies performed on a dedicated device (excluding gamma-camera), to assess the relationship between pretherapeutic SUV and survival at least in univariate analysis, and to have been published as a peer reviewed article in the English, Dutch, or French language. Abstracts were excluded as it cannot be expected to find enough details to assess methodology or survival information to perform meta-analysis. Reports using all modalities of care were included.

Studies were identified by an electronic search on Medline databank using the following keywords: "lung or lung cancer or lung carcinoma or nonsmall cell or non small cell or NSCLC or small cell or SCLC or lung neoplasms" and "positron emission tomography or PET or PET imaging tomography" and "FDG-F^{18}or FDG or ^{18}F-fluorodeoxyglucose or ^{18}F-FDG" and "SUV or standardized uptake value or uptake value" and "prognosis or prognostic factor or outcome," and "survival." The references reported in all the identified studies were used to complete this search which ended in June 2006. The final list of articles eligible for the review was analyzed to identify articles in which there might be overlap in the cohorts of patients used.

Eleven physicians and one biostatistician reviewed all the publications to assess their methodological quality, to determine their eligibility for inclusion in the quantitative meta-analysis and to extract the most important information determining the clinical and PET characteristics. A methodological quality scale was designed for the purpose of this study using the variables available in the publications. This score assessed the clinical and PET reports. The clinical report included the distribution of the expected "prognostic factors" (age, gender, performance status, stage, histology, and weight loss), tumor stage description, staging characteristics (definition of the size of a pathologic mediastinal adenopathy, systematic use of a CT thorax for lung staging, systematic metastatic work-up, systematic use of a CT or magnetic resonance imaging of brain, histologic confirmation of metastatic mediastinal adenopathy, and if the analysis of the relationship between SUV was performed without knowledge of survival results and conversely [double blind]), description of the results of survival analysis (number of patients, number of deaths, follow-up duration, number of patients lost to follow-up, univariate and multivariate analyses, description of statistical tests, survival definition, SUV cutoff definition). The PET report included patients characteristics (weight/height, glycaemia, histologic subtype), ^{18}F-FDG-PET acquisition protocol characteristics (injected dose of ^{18}F-FDG, delay between injection and data acquisition, fasting duration), and technical parameters (investigation area, delay between CT thorax and PET acquisition, SUV formula, type of PET engine, duration of emission time, duration of transmission time, attenuation and reconstruction parameters, type of SUV). The clinical and PET reports were respectively scored on 44 and 40 points. A value between 0 and 2 was attributed to each item. When an item was not applicable to a particular study, it was ruled out. The scores were expressed in percentage of the maximal theoretical value that can be obtained.

The following methodology, as we have already used in previous meta-analyses,[7,8] was applied for aggregating the estimated effects of the ^{18}F-FDG-PET SUV on survival. We measured the impact of SUV on survival by hazard ratio (HR) between the survival distributions of two groups. For each trial, this HR was estimated by a method depending on the results provided in the publication. The most accurate method consisted to retrieve the HR estimate and its variance from the reported results, or to calculate them directly using parameters given by the authors for the univariate analysis: the O-E statistic (difference between numbers of observed and expected events), the confidence interval (CI) for the HR, the logrank statistic, or its p value. If not available, we looked for the total number of events, the number of patients at risk in each group, and the logrank statistic or its p value, allowing calculation of an approximation of the HR estimate. Finally, if the only exploitable data were in the form of graphical representations of the survival distributions, we extracted from them survival rates at some specified times to reconstruct the HR estimate and its variance, with the assumption that the rate of patients censored was constant during the study follow-up.[9] If authors report survival of three or more groups (for example, using several cutoff values for SUV), we pooled the results making a comparison between two groups feasible. The individual HR point estimates were combined after acceptation of the null hypothesis of the homogeneity of the treatment effect across the various trials, using the Peto method[10] to obtain a global HR estimate of the treatment effect. By convention, a HR >1 implied a survival benefit for lower primary tumor SUV. The HR was calculated using a fixed-effects method. In case of significant test for heterogeneity ($p < 0.10$), a random-effects method was applied. This impact of SUV on survival was considered as statistically significant if the 95% CI for the overall HR did not overlap 1. All reported p values were two-tailed.

RESULTS

Sixteen studies, published between 1998 and 2006 were potentially eligible for this review.[11–26] Three studies included similar cohorts of patients[15–17] and we took into account only the most complete one. One study was not assessable for meta-analysis because no quantitative threshold was evaluated (in this study a ^{18}F-FDG-PET scan was interpreted as negative if the tumor uptake of ^{18}F-FDG was less than or equal to mediastinal uptake).[12] In two other studies, patients without definite diagnosis of cancer were reported in the analysis without separated results between cancer patients and the others.[13,14] To reduce the risk of error by including patients without cancer, we performed two meta-analyses one with and one without these two studies.

The principal characteristics of the 13 studies eligible for the meta-analysis are described in Table 1. The majority of the 1474 patients presented with nonmetastatic non-small cell lung cancer (NSCLC); there were only two small cell lung cancers and one carcinoid tumor.[20] Patients were generally staged according to the 1997 edition of the *TNM Classification of Malignant Tumors*. According to the threshold defined by the authors, high SUV was associated with poor prognosis in 11 studies whereas in two, the prognostic role of SUV for survival remained undetermined. The main SUV characteristics reported in the publication are described in Table 2. SUV_{max} was used in eight studies, normalized by body weight in seven (Appendix 1) and by lean body mass in one (Appendix 2). In three studies, the authors used SUV_{mean} normalized by body weight, with different percentages for defining the isocontours of the volume of interest around the tumor. In one study, SUV_{max} and SUV_{mean} normalized by body weight and by lean body mass were studied but only the SUV_{mean} normalized by body weight was used for the survival analysis. The choice of the SUV threshold between patients with high survival and low survival was based on eight cases on a so-called best cutoff, meaning that the authors chose as SUV threshold the value maximizing the logrank test statistic among several survival comparisons. This method is known to lead to a high risk of false-positive result especially if no adjustment of p values for multiplicity is done. In the other publications, the threshold was arbitrarily chosen ($n = 1$), based on the median ($n = 3$) SUV values or was a choice done as validation of results from another author ($n = 1$).

The methodological quality of the studies was moderate. Overall, the median quality score was 57%, ranging from 27 to 68%. The respective median values for the clinical and PET reports were 61% (range 34–80%) and 53% (range 5–65%).

In a first meta-analysis, we excluded the two studies with patients without definite diagnosis of cancer.[13,14] Eleven studies were thus included in the quantitative meta-analysis. The number of patients ranged from 38 to 162 per study, for a total of 1108. The results are detailed in Table 3 and Figure 1. The combined HR for the 11 studies was 2.07 (95% CI: 1.66–2.58) with a fixed-effects model, meaning that high primary tumor SUV was associated with reduced survival. We observed a significant heterogeneity ($p = 0.05$) essentially because of one study.[23] After exclusion of that publication, the test was no more statistically significant ($p = 0.31$). We performed the same analysis using a random-effects model. The HR was 2.13 (95% CI 1.54–2.95). If the two studies[13,14] including patients without histologically proven diagnosis of lung cancer were included in the analysis, we obtained an overall HR of 2.22 (95% CI 1.83–2.70) (fixed effects) or 2.27 (95% CI 1.70–3.02) (random effects, test for heterogeneity $p = 0.06$).

We performed the same analysis excluding the studies proposing a so-called best cutoff. From the eight such studies, two were nevertheless included in the analysis because we were able to use median SUV values instead of the value proposed by the authors.[16,20] The test for heterogeneity was statistically significant ($p = 0.05$). Using a random-effects model, the combined HR was 1.77 (95% CI 1.01–3.12)

TABLE 1. Principal Characteristics of the 13 Studies Included in the Meta-Analysis

Study	Publication Date	*n* Pts	ISS	Stage	Histology	High SUV as Prognostic Factor for Survival
Ahuja et al.[26]	1998	155	1997	I–IV	NSCLC	Unfavorable
Sugawara et al.[23]	1999	38	1986	I–IV	NSCLC	Undetermined
Vansteenkiste et al.[22]	1999	125	1997	I–IIIB	NSCLC	Unfavorable
Dhital et al.[20]	2000	77	1986	≤IIIA	All[a]	Unfavorable
Higashi et al.[16]	2002	57	1997	I–IIIB	NSCLC	Unfavorable
Jeong et al.[18]	2002	73	1997	I–IV	NSCLC	Unfavorable
Downey et al.[25]	2004	100	1997	<IV	NSCLC	Unfavorable
Port et al.[11]	2005	64	1997?	?	NSCLC	Undetermined
Sasaki et al.[24]	2005	162	1997	I–IIIB	NSCLC	Unfavorable
Prevost et al.[21]	2006	120	1997	I–IV?	NSCLC	Unfavorable
Eschmann et al.[19]	2006	137	1997?	III	NSCLC	Unfavorable
Borst et al.[14]	2005	51	?	I–III	NSCLC	Unfavorable
Cerfolio et al.[13]	2005	315	1997	I–IV	NSCLC	Unfavorable

n Pts, number of patients; ISS, date of International Staging System applied in the study; NSCLC, non-small cell lung cancer; SUV, standardized uptake value.

[a] Only 2 small cell lung cancers and 1 carcinoid tumor among 77 patients.

TABLE 2. Main SUV Characteristics Reported in the 13 Publications Assessable for Meta-Analysis

Study	Type of SUV	Correction of SUV	SUV Threshold Definition	SUV Threshold
Ahuja et al.[26]	SUV mean (SUR)	Weight	Best cut-off	10
Sugawara et al.[3]	SUV max	Lean body mass	Median	8.7
Vansteenkiste et al.[22]	SUV max	Weight	Best cut-off	7
Dhital et al.[20]	SUV max	Weight	Best cut-off	15 or 20
Higashi et al.[16]	SUV mean	Weight	Best cut-off	5
Jeong et al.[18]	SUV max	Weight	Best cut-off	7
Downey et al.[25]	SUV max	Weight	Median	9
Port et al.[11]	Unspecified SUV	—	Arbitrary	2.5
Sasaki et al.[24]	SUV max	Weight	Best cut-off	5
Prevost et al.[21]	SUV mean SUV max	Weight Lean body mass	Literature value	10
Eschmann et al.[19]	SUV mean	Weight	Best cut-off	12
Borst et al.[14]	SUV max	Weight	Best cut-off	15
Cerfolio et al.[13]	SUV max	Weight	Median	10

Best cut-off, author maximized the logrank test statistic to determine the best cut-off; SUV, standardized uptake value; SUR, standardized uptake ratio.

TABLE 3. Meta-Analysis of Primary Tumor SUV Prognostic Impact on Survival in Lung Cancer (all Studies)

References	Publication Date	*n* Pts	HR	95% CI
Ahuja et al.[26]	1998	155	2.05	1.24–3.37
Dhital et al.[20]	2000	77	1.30	0.70–2.60
Downey et al.[25]	2004	100	2.60	1.02–6.64
Eschmann et al.[19]	2006	137	1.71	1.00–2.93
Higashi et al.[16]	2002	57	6.20	1.34–28.75
Jeong et al.[18]	2002	73	4.33	1.80–10.45
Port et al.[11]	2005	64	2.36	0.24–22.88
Prevost et al.[21]	2005	120	2.36	1.34–4.15
Sasaki et al.[24]	2005	162	7.66	1.41–41.50
Sugawara et al.[23]	1999	38	0.56	0.21–1.44
Vansteenkiste et al.[22]	1999	125	2.72	1.50–4.94
Borst et al.[14]	2005	51	3.15	1.59–6.22
Cerfolio et al.[13]	2005	315	2.65	1.63–4.31
Meta-analysis including only pathologically proven lung cancer (n = 11 studies)				
Fixed effects		1108	2.07	1.66–2.58
Random effects		1108	2.13	1.54–2.95
Meta-analysis including patients without pathologically proven lung cancer (n = 13 studies)				
Fixed effects		1474	2.22	1.83–2.70
Random effects		1474	2.27	1.70–3.02

HR, hazard ratio; CI, confidence interval; SUV, standardized uptake value; *n* Pts, number of patients.

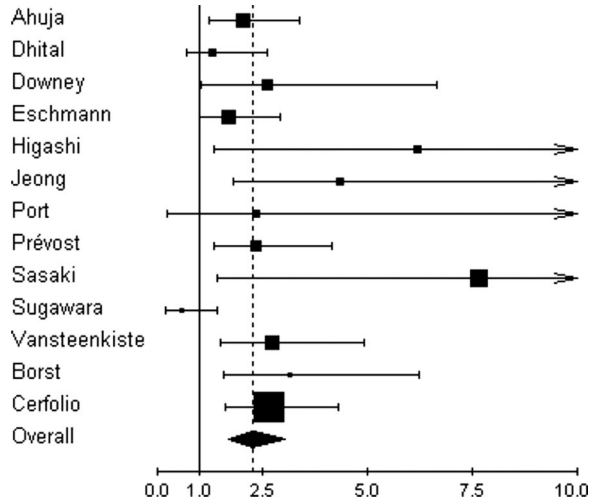

FIGURE 1. Graphical representation of the prognostic role of primary tumor SUV on survival in lung cancer. HR and 95% confidence interval (CI) for survival comparison in studies evaluating primary tumor SUV in lung cancer. HR >1 implied a survival benefit for reduced primary tumor SUVmax. The square size is proportional to the number of patients included in the study. The center of the diamond-shaped lozenge at the bottom of the figure gives the combined HR of the meta-analysis and its extremities the 95% CI HR = 2.27; 95% CI 1.70–3.02 (random-effect model). Total number of patients: 1474. SUV = standardized uptake value.

(Table 4). If the two studies without histologically proven lung cancer[13,14] were included, the combined HR were 2.13, 95% CI: 1.65–2.76 (fixed effects) or 2.08, 95% CI: 1.43–3.04 (random effects). Last, we looked at the role of SUV in nonmetastatic lung cancer, excluding studies incorporating stage IV diseases. Six publications were available, including 658 patients (Table 5). As the test for heterogeneity was not significant ($p = 0.16$), we used a fixed-effects model. The combined HR was 2.09 (95% CI 1.54–2.83). When adding a study without histologically proven lung cancer,[14] the HR was 2.23 (95% CI 1.69–2.95).

DISCUSSION

During the last decade, ^{18}F-FDG-PET has become an important tool used to stage patients with NSCLC. The specific goal of this study was to evaluate its potential as a prognostic marker. The current meta-analysis confirmed that increased SUV of the primary tumor is a poor prognostic factor in patients with NSCLC. It remains unclear, however, if SUV is an independent prognostic feature as compared with stage and performance status.

TABLE 4. Meta-Analysis of Primary Tumor SUV Prognostic Impact on Survival in Lung Cancer (Excluding Studies Proposing a So-Called Best Cutoff)

References	Publication Date	*n* Pts	HR	95% CI
Port et al.[11]	2005	64	2.36	0.24–22.88
Dhital et al.[20]	2000	77	1.30	0.70–2.60
Downey et al.[25]	2004	100	2.60	1.02–6.64
Higashi et al.[16]	2002	57	6.20	1.34–28.75
Prevost et al.[21]	2005	120	2.36	1.34–4.15
Sugawara et al.[23]	1999	38	0.56	0.21–1.44
Fixed effects		456	1.74	1.23–2.44
Random effects		456	1.77	1.01–3.12

HR, hazard ratio; CI, confidence interval; SUV, standardized uptake value; *n* Pts, number of patients.

TABLE 5. Meta-Analysis of Primary Tumor SUV Prognostic Impact on Survival in Lung Cancer (Excluding Studies with Stage IV Diseases)

References	Publication Date	*n* Pts	HR	95% CI
Dhital et al.[20]	2000	77	1.30	0.70–2.60
Eschmann et al.[19]	2006	137	1.71	1.00–2.93
Higashi et al.[16]	2002	57	6.20	1.34–28.75
Sasaki et al.[24]	2005	162	7.66	1.41–41.51
Vansteenkiste et al.[22]	1999	125	2.72	1.50–4.94
Downey et al.[25]	2004	100	2.60	1.02–6.64
Fixed effects		658	2.09	1.54–2.83
Random effects		658	2.27	1.45–3.54

HR, hazard ratio; CI, confidence interval; SUV, standardized uptake value; *n* Pts, number of patients.

Overall, the prognosis of patients with NSCLC is poor, with less than 15% surviving beyond 5 years.[4] Stage, currently determined by the 1997 International Staging System classification, is the most important prognostic factor in NSCLC patients having direct implications in the choice of therapeutic options. As pointed by different authors,[27,28] this staging system still needs to be improved. The IASLC developed a task force to propose revisions for the 7th edition of the *TNM Classification of Malignant Tumors* concerning lung cancer. In this setting, a working group is assessing the potential usefulness of new prognostic factors.

Recently, attention has focused on new biologic factors, and the methodological group of the European Lung Cancer Working Party and others have evaluated the most relevant variables and found that some features could be of interest.[8,29–36] Unfortunately, there are currently no data to support that any of these markers, including genomic profiles, can accurately predict outcomes in NSCLC. The observed HR in these studies was usually smaller in magnitude than those found in the present meta-analysis for primary tumor SUV.

SUV is a semiquantitative index that characterizes the tracer uptake, hence approximating the glucose metabolic rate (Appendix 1). However, SUV estimates suffer from poor reproducibility between centers because of the lack of standardization of the acquisition and processing protocols leading to its assessment (Appendix 2). In our study, this poor reproducibility was evidenced by the broad range of threshold values that have been used in the literature to distinguish between patients with low and high survival (thresholds varying from 2.5 to 20). Despite this variability, we were able to show that SUV was correlated with patient survival. Indeed, our study design calculated an HR for each study center, based on the SUV threshold used in that study, which somehow cancelled the threshold factor. By doing so, we could demonstrate that SUV was certainly worth considering as a prognostic factor, especially as, unlike immunohistochemistry, it can be estimated even when no surgical specimen is present.

To be a practical prognostic factor in routine practice, a single SUV threshold allowing distinguishing between long and short survival patients should be agreed on, or the methodology to be used to determine the optimal threshold for each center should be established. To set a consistent threshold, most sources of variability impacting the SUV estimates (Appendix 2) should be removed or at least controlled, for example, by phantom calibration.[37] Reducing the large variability currently affecting SUV estimates would probably enhance the prognostic value of SUV. In our meta-analysis, we could not take into account the variable conditions in which the SUV were obtained given the poor quality scores of the PET reports. A meta-analysis considering the individual patient data (IPD) will be needed to try to compensate for the large heterogeneity of the reported SUV.

Some biases might have occurred in our analysis. Indeed, some studies were not included in our meta-analysis because, e.g., separate data for lung cancer patients were not available. We did not look at trials presented only on their abstract form or at unpublished studies. Thus, some studies might have not been taken into account. We limited this problem by discussing with experts in the field during regular meetings of the IASLC International Staging Committee. We carefully looked at the possibility of patients' duplication by reporting the same cohorts in different publications. This led us to suppress two articles,[15,17] although no reference to such duplicates was reported by the authors. Some difficulties could happen when analyzing and comparing the results of the individual trials. The stage and treatment case mix were somewhat different through the studies; in particular few treatment results were reported by the authors.

To avoid some biases of a literature-based meta-analysis, we aim to confirm our results in a meta-analysis of IPD. Literature-based meta-analyses has the advantage of including published trials immediately available for analysis and which results can be checked by everyone. Although we found in previous publications similar results with literature-based as with IPD meta-analyses,[38,39] IPD add some interest like incorporating unpublished trials, updating results (allowing to have longer survival follow-up), and particularly allowing for multivariate analyses, adjusting for other variables, and subgroup analyses.

In conclusion, metabolic activity of primary tumor, reflected by SUV measurement with ^{18}F-FDG-PET scan is a prognostic factor in patients with NSCLC. It still needs to be

compared with stage and performance status in a formal analysis to determine if this adds prognostic value. We are currently planning a meta-analysis based on IPD that will allow multivariate analysis and potentially reduce biases related to literature-based meta-analyses. These results may be of particular importance in the view of the forthcoming, 7th edition of the *TNM Classification for Lung Cancer*.

REFERENCES

1. Graziano SL. Non-small cell lung cancer: clinical value of new biological predictors. *Lung Cancer* 1997;17(Suppl 1):S37–S58.
2. Kanters SD, Lammers JW, Voest EE. Molecular and biological factors in the prognosis of non-small cell lung cancer. *Eur Respir J* 1995;8: 1389–1397.
3. Sculier JP, Paesmans M, Libert P, et al. Facteurs pronostiques dans le cancer bronchique non à petites cellules avancé: expériences de l'European Lung Cancer Working Party. *Rev Mal Respir* 1997;14:445–449.
4. Mountain CF. Revisions in the international system for staging lung cancer. *Chest* 1997;111:1710–1717.
5. Gould MK, Maclean CC, Kuschner WG, Rydzak CE, Owens DK. Accuracy of positron emission tomography for diagnosis of pulmonary nodules and mass lesions: a meta-analysis. *JAMA* 2001;285:914–924.
6. 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.
7. Berghmans T, Paesmans M, Meert AP, et al. Survival improvement in resectable non-small cell lung cancer with (neo)adjuvant chemotherapy: results of a meta-analysis of the literature. *Lung Cancer* 2005;49:13–23.
8. Meert AP, Martin B, Delmotte P, et al. The role of EGF-R expression on patient survival in lung cancer: a systematic review with meta-analysis. *Eur Respir J* 2002;20:975–981.
9. Parmar MK, Torri V, Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. *Stat Med* 1998;17:2815–2834.
10. Yusuf S, Peto R, Lewis J, Collins R, Sleight P. Beta blockade during and after myocardial infarction: an overview of the randomized trials. *Prog Cardiovasc Dis* 1985;27:335–371.
11. 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.
12. Marom EM, Sarvis S, Herndon JE, Patz EF Jr. T1 lung cancers: sensitivity of diagnosis with fluorodeoxyglucose PET. *Radiology* 2002; 223:453–459.
13. 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. *J Thorac Cardiovasc Surg* 2005;130:151–159.
14. 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.
15. Guo J, Higashi K, Yokota H, et al. In vitro proton magnetic resonance spectroscopic lactate and choline measurements, 18F-FDG uptake, and prognosis in patients with lung adenocarcinoma. *J Nucl Med.* 2004;45: 1334–1339.
16. 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.
17. Higashi K, Ueda Y, Ayabe K, et al. FDG PET in the evaluation of the aggressiveness of pulmonary adenocarcinoma: correlation with histopathological features. *Nucl Med Commun* 2000;21:707–714.
18. Jeong HJ, Min JJ, Park JM, et al. Determination of the prognostic value of [(18)F]fluorodeoxyglucose uptake by using positron emission tomography in patients with non-small cell lung cancer. *Nucl Med Commun* 2002;23:865–870.
19. Eschmann SM, Friedel G, Paulsen F, et al. Is standardised (18)F-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.
20. Dhital K, Saunders CA, d PT, et al. [(18)F]Fluorodeoxyglucose positron emission tomography and its prognostic value in lung cancer. *Eur J Cardiothorac Surg* 2000;18:425–428.
21. Prevost S, Boucher L, Larivee P, Boileau R, Benard F. 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. Vansteenkiste JF, Stroobants SG, Dupont PJ, et al. Prognostic importance of the standardized uptake value on (18)F-fluoro-2-deoxy-glucose-positron emission tomography scan in non-small-cell lung cancer: an analysis of 125 cases. Leuven Lung Cancer Group. *J Clin Oncol* 1999;17:3201–3206.
23. Sugawara Y, Quint L, Iannettoni M, 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.
24. 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.
25. 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.
26. Ahuja V, Coleman RE, Herndon J, Patz EF Jr. The prognostic significance of fluorodeoxyglucose positron emission tomography imaging for patients with nonsmall cell lung carcinoma. *Cancer* 1998;83:918–924.
27. Leong SS, Rocha Lima CM, Sherman CA, Green MR. The 1997 International staging system for non-small cell lung cancer: have all the issues been addressed? *Chest* 1999;115:242–248.
28. Berghmans T, Lafitte JJ, Thiriaux J, et al. Survival is better predicted with a new classification of stage III unresectable non-small cell lung carcinoma treated by chemotherapy and radiotherapy. *Lung Cancer* 2004;45:339–348.
29. Berghmans T, Paesmans M, Mascaux C, et al. Thyroid transcription factor 1—a new prognostic factor in lung cancer: a meta-analysis. *Ann Oncol* 2006;17:1673–1676.
30. Mascaux C, Iannino N, Martin B, et al. The role of RAS oncogene in survival of patients with lung cancer: a systematic review of the literature with meta-analysis. *Br J Cancer* 2005;92:131–139.
31. Choma D, Daures JP, Quantin X, et al. Aneuploidy and prognosis of non-small-cell lung cancer: a meta-analysis of published data. *Br J Cancer* 2001;85:14–22.
32. Meert AP, Paesmans M, Martin B, et al. The role of microvessel density on the survival of patients with lung cancer: a systematic review of the literature with meta-analysis. *Br J Cancer* 2002;87:694–701.
33. Steels E, Paesmans M, Berghmans T, et al. Role of p53 as a prognostic factor for survival in lung cancer: a systematic review of the literature with a meta-analysis. *Eur Respir J* 2001;18:705–719.
34. Meert AP, Martin B, Paesmans M, et al. The role of HER-2/neu expression on the survival of patients with lung cancer: a systematic review of the literature. *Br J Cancer* 2003;89:959–965.
35. Martin B, Paesmans M, Berghmans T, et al. Role of Bcl-2 as a prognostic factor for survival in lung cancer: a systematic review of the literature with meta-analysis. *Br J Cancer* 2003;89:55–64.
36. Delmotte P, Martin B, Paesmans M, et al. VEGF et survie des patients avec cancer pulmonaire: une revue systématique de la littérature et méta-analyse. *Rev Mal Respir* 2002;19:577–584.
37. van Baardwijk A, Dooms C, van Suylen RJ, et al. The maximum uptake of (18)F-deoxyglucose on positron emission tomography scan correlates with survival, hypoxia inducible factor-1alpha and GLUT-1 in non-small cell lung cancer. *Eur J Cancer* 2007;43:1392–1398.
38. Auperin A, Arriagada R, Pignon JP, et al. Prophylactic cranial irradiation for patients with small-cell lung cancer in complete remission. Prophylactic Cranial Irradiation Overview Collaborative Group. *N Engl J Med* 1999;341:476–484.
39. Meert AP, Paesmans M, Berghmans T, et al. Prophylactic cranial irradiation in small cell lung cancer: a systematic review of the literature with meta-analysis. *BMC Cancer* 2001;1:5.
40. Bacharach SL, Sundaram SK. 18F-FDG in cardiology and oncology: the bitter with the sweet. *J Nucl Med* 2002;43:1542–1544.
41. Soret M, Bacharach SL, Buvat I. Partial-volume effect in PET tumor imaging. *J Nucl Med* 2007;48:932–945.
42. Boellaard R, Krak NC, Hoekstra OS, Lammertsma AA. Effects of noise, image resolution, and ROI definition on the accuracy of standard uptake values: a simulation study. *J Nucl Med* 2004;45:1519–1527.

43. Krak NC, Boellaard R, Hoekstra OS, Twisk JW, Hoekstra CJ, Lammertsma AA. Effects of ROI definition and reconstruction method on quantitative outcome and applicability in a response monitoring trial. *Eur J Nucl Med Mol Imaging* 2005;32:294–301.
44. Kim CK, Gupta NC, Chandramouli B, Alavi A. Standardized uptake values of FDG: body surface area correction is preferable to body weight correction. *J Nucl Med* 1994;35:164–167.
45. Langen KJ, Braun U, Rota KE, et al. The influence of plasma glucose levels on fluorine-18-fluorodeoxyglucose uptake in bronchial carcinomas. *J Nucl Med* 1993;34:355–359.
46. Zhuang H, Pourdehnad M, Lambright ES, et al. Dual time point 18F-FDG PET imaging for differentiating malignant from inflammatory processes. *J Nucl Med* 2001;42:1412–1417.

APPENDIX 1: DEFINITIONS OF SUV_{MAX} AND SUV_{MEAN}

The standardized uptake value (SUV) is usually defined as the tracer uptake in the tumor divided by the injected dose normalized by the patient weight:

$$\text{Suv} = \text{Tumor Uptake (kBq/mL)}/[\text{injected Dose (kBq)}/\text{Patient Weight (g)}] \quad (1)$$

Assuming the patient has a density of 1 (1 ml = 1 g), the SUV is a dimension less index. If the tracer was uniformly distributed throughout the body, the SUV would be 1 in any organ. As tumors usually have an enhanced metabolic rate, most tumors show an elevated SUV in ^{18}F-FDG PET (SUV >1). However, SUV is only an approximate indicator of the glucose metabolic rate, and is by no mean an accurate measurement of this rate.[40]

Although SUV is a widely accepted index for assessing tumor uptake in ^{18}F-FDG PET, there is no consensus regarding how to estimate it. The major differences between the estimation methods come from the way tumor uptake is measured.[41] Two classic measurement methods are:

Considering the tumor uptake is given by the maximum pixel value in the tumor, which yields SUV_{max}:

$$\text{Suvmax} = \text{Max Pixel Value In The Tumor (kBq/mL)}/[\text{injected Dose (kBq)}/\text{Patient Weight (g)}] \quad (2)$$

Measuring the tumor uptake as the mean pixel value in a volume of interest (VOI) around the tumor, which yields SUV_{mean}.

$$\text{Suvmax} = \text{Mean Pixel Value In A Tumor Voi (kBq/mL)}/[\text{injected Dose (kBq)}/\text{Patient Weight (g)}] \quad (3)$$

Unlike SUV_{max}, SUV_{mean} depends on the way the VOI around the tumor is drawn. To avoid manual drawing that makes the result dependent on the operator, an isocontour, defined as a percentage of the maximum pixel value in the tumor (typically between 50 and 80%) is often used. However, there is no consensus on the percentage that should be used, so even the way SUV_{mean} is calculated can greatly vary among centers. In addition, there are some other ways to define the VOI, such as using a fixed sized region regardless of the tumor size.

APPENDIX 2: FACTORS INFLUENCING THE RESULTS OF THE SUV CALCULATION

The most important sources of variability in SUV estimates are listed below:

- The method of tumor uptake estimate (Appendix 1): it greatly impacts the SUV, which can vary by a factor of about 2 depending on the VOI considered to measure the tumor uptake.[41–43]
- The spatial resolution in the reconstructed images (depending itself on the image reconstruction algorithm and of its parameters), which directly affects the blurring of the tumor in the images, hence the tumor pixel values. The same tumor n with two PET imaging systems with different spatial resolutions will appear to have different uptakes through the partial volume effect.[41,42] This is especially true for SUV_{max}, which can vary by more than 10% depending on the spatial resolution in the reconstructed images.
- The way the PET images have been compensated for physical biases such as attenuation. For instance, CT-based attenuation correction tends to yield higher SUV (from 10 to 50%) than attenuation correction based on a conventional PET transmission device.
- The normalization factor used to estimate the FDG made available to the tumor (denominator of Eq. 1). Most often, the injected activity is normalized by the patient weight (Eq. 1), but other normalizations, such as using the lean body mass or the body surface area[44] have been proposed to account for the fact that body fat does not have the same FDG uptake as lean tissues.
- The plasma glucose level of the patient, as the FDG competes with the plasma glucose: lower SUV are observed in fed patients compared with patients under fasting conditions.[45] Normalization to account for the plasma glucose level has been proposed.
- The delay between the injection time and the imaging time: the longer this delay, the higher the SUV, as equilibrium is usually not reached at 45 to 60 minutes postinjection.[46] Changes in SUV between 45 minutes and 90 minutes postinjection scans can be of about 20%.

As new corrections become available, SUV will also depend on:

- Whether the PET images are compensated for respiratory motion, e.g., using respiratory gating, as respiratory motion introduces blur in the images, hence contributes to lowering the pixel values in the tumor.
- Whether the PET images are compensated for partial volume effects.