Methods to Monitor Response to Chemotherapy in Non–Small Cell Lung Cancer with ^{18}F-FDG PET

Corneline J. Hoekstra, MD[1,2]; Otto S. Hoekstra, MD, PhD[1,3]; Sigrid G. Stroobants, MD[4]; Johan Vansteenkiste, MD, PhD[5]; Johan Nuyts, PhD[4]; Egbert F. Smit, MD, PhD[2]; Maarten Boers, MD, PhD[3]; Jos W.R. Twisk, PhD[3]; and Adriaan A. Lammertsma, PhD[1]

[1]Clinical PET Centre, VU University Medical Centre, Amsterdam, The Netherlands; [2]Department of Pulmonary Medicine, VU University Medical Centre, Amsterdam, The Netherlands; [3]Department of Clinical Epidemiology and Biostatistics, VU University Medical Centre, Amsterdam, The Netherlands; [4]Department of Nuclear Medicine, University Hospital Gasthuisberg, Leuven, Belgium; and [5]Department of Pulmonary Medicine, University Hospital Gasthuisberg, Leuven, Belgium

PET using ^{18}F-FDG is a promising technique to monitor response in oncology. Unfortunately, a multitude of analytic methods is in use. To date, it is not clear whether simplified methods could replace complex quantitative methods in routine clinical practice. The aim of this study was to select those methods that would qualify for further assessment in a future prospective response-monitoring study by comparing results with patient outcome. **Methods:** Dynamic ^{18}F-FDG PET scans were obtained on 2 groups of patients. First, 10 patients with advanced non–small cell lung cancer (NSCLC) were scanned on consecutive days before treatment to assess test–retest variability. Second, 30 scans were obtained on 19 patients with locally advanced NSCLC as part of an ongoing response-monitoring study. These scans were analyzed by 2 observers to assess observer variability. In addition, these studies were used to compare various methods with the gold standard, full kinetic analysis (nonlinear regression [NLR]). **Results:** Using an image-derived input function, NLR showed excellent test–retest and observer agreement confirming that it could be used as a gold standard method. From a total of 34 analytic methods, 10 showed good correlation with NLR. Taking into account the degree of complexity of the methods, 4 remain for further evaluation. **Conclusion:** The optimal method for analysis of ^{18}F-FDG PET data was determined for several levels of complexity. Four methods need to be evaluated further to determine the optimal trade-off between simplicity and accuracy for routine clinical practice.

Key Words: response monitoring; ^{18}F-FDG PET; cancer; methodology

J Nucl Med 2002; 43:1304–1309

At present, combined modality treatment (chemotherapy followed by surgery or radiotherapy) for patients with locally advanced non–small cell lung cancer (NSCLC stage IIIA) is being studied extensively. It is clear, however, that a substantial number of patients do not benefit from such intensive treatment. For example, distant metastases frequently appear during or shortly after induction chemotherapy (*1*), testifying to the inaccuracy of current staging algorithms. Conventional techniques used to monitor therapeutic effects in oncology, such as CT and MRI, are based on morphologic changes and show limited accuracy (*2*). It is expected that functional changes will precede morphologic changes and, therefore, techniques that image function rather than anatomy might provide better accuracy to monitor response and overall treatment outcome.

The value of ^{18}F-FDG PET as such a functional imaging technique for monitoring response is still under investigation (*3,4*). Initial studies have suggested additional value for patient management (*3–6*). Unfortunately, a full assessment of the value of ^{18}F-FDG PET for monitoring response to therapy is complicated by the presence of a variety of scanning protocols and analytic techniques. To date, it is not clear which analytic method is the most accurate (*7*). Nonlinear regression (NLR) is generally accepted as the gold standard method. However, it is unclear whether it still can be considered as such when an image-derived input function is used.

At present, no studies in large groups of NSCLC patients have been performed to assess the actual value of ^{18}F-FDG PET as a response-monitoring technique. Moreover, the best implementation still has to be defined. In preparation of a large multicenter trial on the value of ^{18}F-FDG PET for monitoring response in patients with NSCLC, this study was performed to assess (a) whether the reliability of full compartmental analysis with NLR, being the most quantitative method, accords with its status as the gold standard

Received Feb. 6, 2002; revision accepted Jun. 11, 2002.

For correspondence or reprints contact: Adriaan A. Lammertsma, PhD, Clinical PET Centre, VU University Medical Centre, P.O. Box 7057, 1007 MB Amsterdam, The Netherlands.

E-mail: aa.lammertsma@vumc.nl

when an image-derived input function is used; and (b) the degree of correspondence of different (simplified) analytic methods with the gold standard.

The design of this study is in line with the European Organization for Research and Treatment of Cancer (EORTC) recommendations (*8*) for monitoring response using ^{18}F-FDG PET. As such, this study could serve as a model for further response-monitoring studies on other tumor types and also for the evaluation of other tracers.

MATERIALS AND METHODS

Scans were performed on 2 separate groups of patients. First, a group of 10 patients (8 men, 2 women; mean age, 53 ± 6.8 y) with NSCLC stage IIIB/IV was scanned twice on consecutive days before the start of chemotherapy to assess test–retest variability. Second, in a separate group, 30 randomly selected dynamic scans were used, which were obtained on 19 patients (15 men, 4 women; mean age, 59.4 ± 8.1 y) with NSCLC stage IIIA-N2, as part of an ongoing response-monitoring study.

Scans were performed using a state-of-the-art PET scanner (ECAT EXACT HR+; Siemens/CTI, Knoxville, TN). This scanner has an axial field of view of 15 cm, divided into 63 contiguous planes. The patient was positioned supine on the scanner bed with the tumor in the center of the axial field of view.

All patients fasted for 6 h before scanning. Patients received 2 venous catheters: one for injection of ^{18}F-FDG contralateral to the tumor and the other for venous blood sampling. Acquisition started with a 10- to 15-min transmission scan to correct for photon attenuation (*9*), followed by a bolus injection of 370 MBq ^{18}F-FDG in 5 mL saline through an injector (Medrad International, Maastricht, The Netherlands) at 0.8 mL/s, after which the line was flushed with 42 mL saline (2.0 mL/s). Simultaneous with the injection of ^{18}F-FDG, a dynamic emission scan (in 2-dimensional mode) was started with a total duration of 60 min with variable frame length (6 × 5 s, 6 × 10 s, 3 × 20 s, 5 × 30 s, 5 × 60 s, 8 × 150 s, and 6 × 300 s). All dynamic scan data were corrected for dead time, decay, scatter, randoms, and photon attenuation and were reconstructed as 128 × 128 matrices using filtered backprojection (FBP) with a Hanning filter (cutoff, 0.5 cycle/pixel). This resulted in a transaxial spatial resolution of around 7-mm full width at half maximum. Before injection of ^{18}F-FDG, a blood sample was collected for determination of the plasma glucose level. In addition, 3 venous blood samples were drawn at 35, 45, and 55 min after ^{18}F-FDG injection as quality control for the image-derived input function (*10,11*) and for plasma glucose measurement (hexokinase method, Hitachi 747; Boehringer Mannheim, Mannheim, Germany).

Data Analysis

Three-dimensional regions of interest (ROIs) were defined semiautomatically over the tumor using a threshold of 50% of the maximum pixel value within the tumor. For this purpose, the last 3 frames of the sinograms were summed and reconstructed using ordered-subset expectation maximization with 2 iterations and 12 subsets followed by postsmoothing of the reconstructed image using a 5-mm full width at half maximum gaussian filter to obtain the same resolution as FBP data (*12*). For each tumor ROI, a mirror region was defined by copying the tumor ROI to the contralateral healthy lung. In addition, using FBP data, ROIs were defined manually over the aortic arch, left ventricle, and left atrium to obtain an image-derived input function, as described (*10*).

Test–retest variability was assessed by analyzing the 2 sets of 10 scans from the first group of patients by a single observer. For intraobserver variability, the 30 scans of the second group of patients were analyzed twice by 1 observer on different days (maximum interval, 1–7 d). To assess interobserver variability, the same 30 scans were analyzed by an independent second observer using the same method and software for analysis.

Analysis was performed using the following methods: ratio of tumor to normal tissue (T/N), standardized uptake value (SUV) for the intervals 40–60 and 50–60 min after injection (with several correction factors), NLR using the standard 2-tissue compartment model with 3 (3k) and 4 (4k) rate constants, a blood volume component and an image-derived input function, and the Patlak graphical analysis (*13*). Furthermore, the 2-ROI, 6-parameter model by Wu et al. (*14*), the correlation coefficient filtered influx constant image (correlative imaging) by Zasadny et al. (*15*), the total lesion evaluation (TLE) method by Wu et al. (*16*), the simplified kinetic method (SKM) described by Hunter et al. (*17*) for the intervals 40–60 and 50–60 min after injection, and the net influx constant by Sadato et al. (*18*) were investigated. All methods were applied to the same dataset. A more detailed description of the various methods can be found elsewhere (*7*).

In this study the lumped constant used was set to 1 and was assumed to be constant over time because no studies on the actual value of the lumped constant in tumors outside the central nervous system have been reported.

Statistics

The presence of a fourth rate constant ([k_4] i.e., dephosphorylation of FDG-6-PO_4 back to FDG) and the need to include this in the model was assessed by comparing residual sum of squares of fits with and without a k_4 parameter using Akaike (*19*) and Schwarz (*20*) criteria.

The level of agreement within a method was assessed with intraclass correlation coefficients (ICCs) (*21*). In this setting, the ICC is the proportion of total variance that can be ascribed to true differences within the method. Values for the ICC range from 0 to 1. Values close to 0 indicate poor agreement between repeated observations (i.e., most of the variance is due to measurement error); values close to 1 indicate high agreement. ICCs were assessed for NLR using test–retest and inter- and intraobserver data to establish whether NLR could be used as a gold standard method. All measures of glucose consumption (obtained with the analytic methods mentioned above) were compared with those of NLR using regression analysis (Pearson). A cutoff value for r^2 of 0.95 was used to select these methods with results closest to those obtained with NLR.

Regression analysis determined slopes and intercepts of the correlation between NLR and other methods. These results can be used to draw nomograms that allow approximation of NLR values.

RESULTS

The 3k model provided significantly better fits than the 4k model in 26 (87%) and 27 (90%) of 30 scans according to Akaike (*19*) and Schwarz (*20*) criteria, respectively. In other

words, the data did not support inclusion of a fourth rate constant and, consequently, it was set to zero in this study.

The mean value ± SD for the plasma glucose level found in this study was 5.3 ± 0.6 mmol/L (range, 3.9–6.5 mmol/L).

NLR using the standard (3k) 2-tissue compartment model with a blood volume component and using an image-derived input function proved to have an excellent test–retest variability (ICC, 0.95; 95% confidence interval [CI], 0.82–0.99). The inter- and intraobserver variabilities were also very good (ICC, 0.98; 95% CI, 0.96–0.99) for both. It was concluded that NLR using the above-mentioned model, with an image-derived input function, could indeed be used as a gold standard. The mean value ± SD for glucose consumption using NLR with 3 rate constants (3k model) was 0.15 ± 0.08 μmol/mL/min (range, 0.02–0.33 μmol/mL/min).

Of the total of 34 alternative methods investigated, 24 had a suboptimal correlation with NLR ($r^2 < 0.95$; Table 1). The best correlations with NLR were found with the SUV corrected for body surface area (BSA) and plasma glucose (SUV_{BSAg}) measured at 40–60 and 50–60 min, the SKM at 40–60 and 50–60 min, the net influx constant method corrected for BSA, and the Patlak graphical analysis (Table 1; Fig. 1). The best scanning period for Patlak graphical analysis was found to be from 10 to 60 min.

In Table 1 the slope and intercept of the regression lines are given, allowing NLR results to be converted to any of the other methods and vice versa.

TABLE 1
Comparison of Glucose Consumption Measurements with Analytic Methods by Regression Analysis with Those of NLR

Method	r^2	Slope	SE	Intercept	SE
T/N 40–60 min	0.712	58.10	7.25	1.471	1.231
T/N 50–60 min	0.701	61.91	7.94	1.592	1.348
SUV_W, 40–60 min	0.898	30.54	1.94	0.768	0.332
SUV_{Wg}, 40–60 min	0.910	34.77	2.06	0.009	0.356
SUV_{LBM}, 40–60 min	0.845	27.01	2.20	0.987	0.376
SUV_{LBMg}, 40–60 min	0.925	30.42	1.63	0.326	0.279
SUV_{BSA}, 40–60 min	0.914	765.5	44.38	22.54	7.60
SUV_{BSAg}, 40–60 min	0.966*	865.4	30.96	3.813	5.30
SUV_W, 50–60 min	0.898	32.11	2.04	0.753	0.349
SUV_{Wg}, 50–60 min	0.908	36.53	2.19	−0.039	0.376
SUV_{LBM}, 50–60 min	0.846	28.40	2.29	0.986	0.392
SUV_{LBMg}, 50–60 min	0.925	31.97	1.72	0.296	0.294
SUV_{BSA}, 50–60 min	0.914	804.8	46.67	22.30	7.99
SUV_{BSAg}, 50–60 min	0.964*	909.4	33.31	2.767	5.71
Net influx constant (W)	0.910	1.027	0.061	0	0.011
Net influx constant (BSA)	0.966*	0.871	0.031	0.004	0.005
SKM 40–60 min†	0.974*	0.893	0.028	0.007	0.005
SKM 50–60 min†	0.972*	0.891	0.028	0.006	0.005
Patlak (10–60 min) MR_{glu}	0.984*	0.931	0.022	0.004	0.004
Patlak (10–45 min) MR_{glu}	0.982*	0.941	0.023	0.004	0.004
Patlak (10–30 min) MR_{glu}	0.951*	0.959	0.042	0.009	0.007
Patlak (20–60 min) MR_{glu}	0.972*	0.921	0.029	0.008	0.005
Patlak (30–60 min) MR_{glu}	0.968*	0.893	0.031	0.013	0.005
TLE (0.5) MR_{glu}‡	0.638	83.31	11.83	−4.095	2.023
TLE (0.5) MR_{glu}§	0.619	0.318	0.047	0.041	0.008
TLE (0.6) MR_{glu}‡	0.658	75.84	10.33	−4.263	1.769
TLE (0.6) MR_{glu}§	0.679	0.370	0.048	0.048	0.008
TLE (0.7) MR_{glu}‡	0.676	67.04	8.77	−4.402	1.502
TLE (0.7) MR_{glu}§	0.699	0.451	0.056	0.053	0.009
TLE (0.8) MR_{glu}‡	0.679	53.71	6.97	−4.168	1.194
TLE (0.8) MR_{glu}§	0.835	0.577	0.049	0.062	0.008
TLE (0.9) MR_{glu}‡	0.536	29.27	5.14	−2.800	0.881
TLE (0.9) MR_{glu}§	0.815	0.739	0.080	0.075	0.016
2-ROI, 6-parameter model	0.419	2.406	0.556	−0.073	0.094

*$r^2 > 0.95$.
†Simplified kinetic method described by Hunter et al. (*17*).
‡Total lesion evaluation method described by Wu et al. (*14*).
§Correlative imaging method described by Zasadny et al. (*15*).
W = body weight; Wg = W plasma glucose; LBM = lean body mass; LBMg = LBM plasma glucose; BSA = body surface area; BSAg = BSA plasma glucose; MR_{glc} = metabolic rate of glucose.

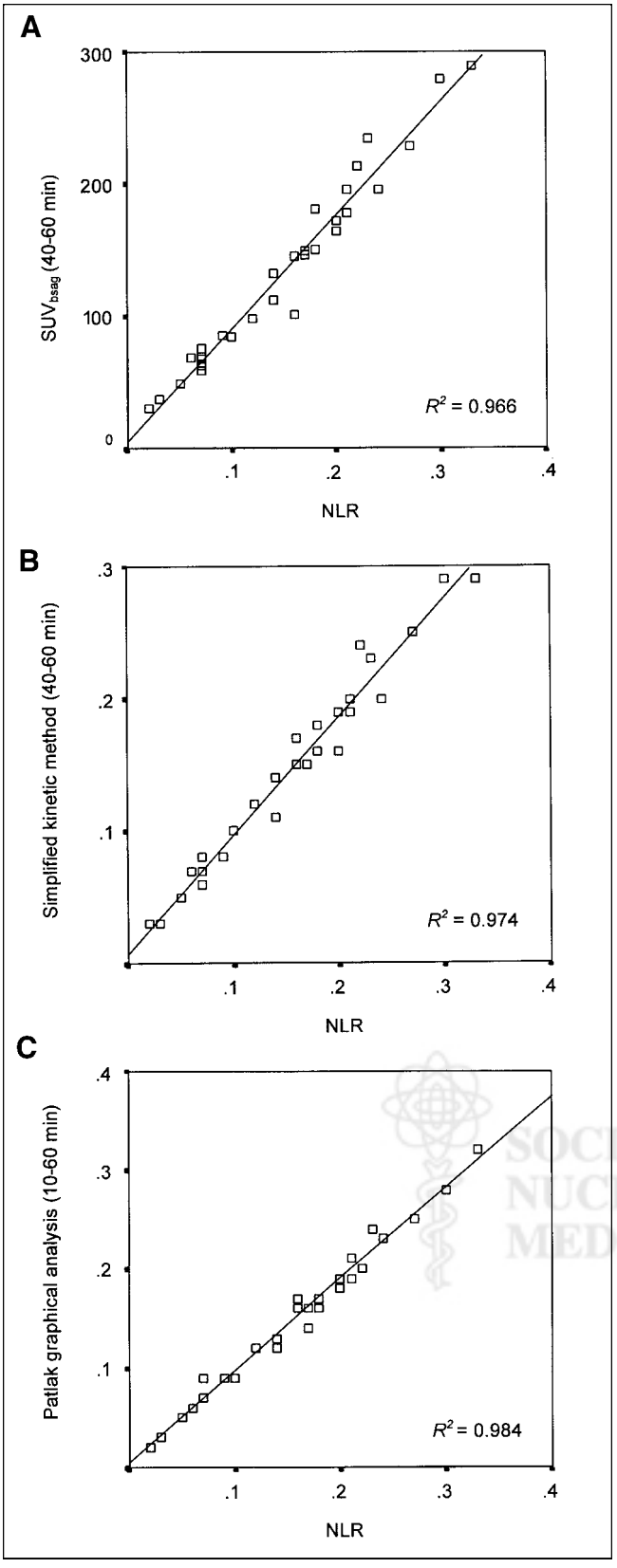

FIGURE 1. Scatter plots of correlation between NLR and SUV_{BSAg} (A), simplified kinetic method (B), and Patlak graphical analysis (C).

DISCUSSION

^{18}F-FDG PET appears to be a promising technique to monitor response to chemotherapy (*8*). Unfortunately, its true value is unknown, as no meta-analysis of reported data is feasible because of the variety of analytic methods used in different studies. Various approaches have been used, ranging from visual assessment (qualitative), through semiquantitative indices, to full kinetic analyses of ^{18}F-FDG uptake (*7*). It is by no means clear whether these methods have similar sensitivity for monitoring changes after therapy. Because interest in response monitoring is growing, selection of the most appropriate method is vital to determine the role of ^{18}F-FDG PET in this field.

For each tumor the presence and possible effect of a fourth rate constant (k_4) should be assessed to make a proper decision on whether to include a k_4 parameter in the model (*22,23*). In this study on NSCLC, a model with 3 rate constants provided significantly better fits than other models that included a fourth rate constant.

Various analytic methods were compared in our study by applying them to the same dataset. One question was whether NLR, generally accepted as the gold standard method, could still be considered as such when an image-derived input function is used. For full kinetic modeling (NLR), both dynamic scanning (tissue time–activity curve) and arterial sampling (arterial plasma time–activity curve) should be used. Arterial cannulation is less suitable for routine clinical response studies, where repetitive scans are required. Recently, measurement of the metabolic rate of glucose (MR_{glu}) using an image-derived input function obtained from thoracic vascular structures has been validated against arterial blood sampling (*11*). In this study an image-derived input function was used, applying quality control measures as defined previously (*10*). However, manual ROI definition of the vascular structures (aorta, atrium, or ventricle) could introduce inter- and intraobserver variability. In addition, noise could be introduced by the limited number of counts acquired in each frame. Nevertheless, the ICC for both test–retest and intra- and interobserver variability was excellent, indicating that even when an image-derived input function is used NLR is a reproducible method for measuring glucose consumption. Therefore, it was used in our study as the gold standard for assessing other simplified methods.

Results of the other (simplified) analytic methods were correlated with NLR (Table 1; Fig. 1) and a cutoff value of $r^2 > 0.95$ for selected methods that approach NLR. Although test–retest and inter- and intraobserver variation could also have been determined for the simplified methods, this analysis did not seem to be very useful. The main variability in NLR will be caused by variation in defining ROI over the blood pool for the image-derived input curve. By combining as many ROIs as possible over 3 different blood-pool structures, these variations can be minimalized

(*10*). For simplified methods such as SUV, only tumor ROIs need to be defined and, as a semiautomatic program is used, variation will be much less than for NLR. Test–retest data only address variability for unchanged conditions. However, uptake of ^{18}F-FDG in the tumor and distribution of ^{18}F-FDG throughout the body may be expected to change because of the effects of therapy. These changes are taken care of in NLR (for each study a new input function is defined) but are not necessarily incorporated in simplified methods (*24,25)*. Therefore, quantitative accuracy (i.e., correspondence with glucose consumption [NLR]) appears to be more relevant for response-monitoring studies.

As expected in the case of no dephosphorylation ($k_4 = 0$), Patlak graphical analysis approaches the accuracy of NLR with the highest correlation obtained for the interval from 10 to 60 min after injection. Although this should be investigated for each tumor type, it is clear that, for NSCLC, Patlak graphical analysis is an excellent alternative for NLR. This would allow for a simpler scanning protocol (less frames) and the option to generate functional glucose consumption (MR_{glu}) images.

For routine clinical use, an even simpler and shorter scanning protocol (single frame) would be preferable. However, methods using such a protocol have been criticized because they are based on several simplifying assumptions, which may not be valid in repeated (response monitoring) studies and thus may introduce errors in the final analysis (*7*). As for the semiquantitative SUV, correcting for BSA and plasma glucose appeared to be preferable over correcting for body weight or lean body mass in agreement with previous studies (*25–27*). Our findings contradict the recommendations made by the EORTC PET study group (*8*). They did not recommend correction for plasma glucose because of concern about the accuracy of measurements in many institutes. Our study indicates that an accurate correction should be used.

The net influx constant method described by Sadato et al. (*18*) also showed a good correlation with NLR. However, this method is directly proportional to SUV because it uses a fixed scaling factor for translating SUV into a measure of MR_{glu}. The simplified kinetic method of Hunter et al. (*17*) showed even better agreement with NLR than with SUV. The method holds promise for routine clinical applications because it is requires only a static scan and a venous blood sample for the tail of the plasma input function (without having to measure the plasma curve itself).

The analytic methods used in ^{18}F-FDG PET studies are simplifications of the underlying physiology. Variability in results can be induced by many factors, most of which can potentially be avoided at the cost of more complicated study protocols. Simplifications in the models can be implemented, however, at the cost of accuracy. This is important to keep in mind when selecting an analytic method for a study protocol. Selection of the degree of accuracy needed (i.e., method required) will depend on how large the differences are between responding and nonresponding tumors.

It is not possible to select the optimal method for response monitoring in NSCLC using the results of our study. This will be possible only in a larger series of patients where response data are compared with clinical outcome because selection of the optimal trade-off between accuracy and simplicity will depend on the actual changes being measured. The purpose of our study was primarily to reduce the multitude of available methods to a limited number of potentially worthwhile techniques, each with a different degree of complexity regarding correction factors used and scanning protocols. We suggest that the following methods be compared with NLR in a prospective clinical study: Patlak graphical analysis from 10 to 60 min, the simplified kinetic method at 40–60 min, and SUV corrected for BSA and plasma glucose at 40–60 min. Such a study is currently in progress in our institutions.

CONCLUSION

Even with an image-derived input function, NLR has excellent test–retest and observer agreement supporting its use as a gold standard method. Of a total of 34 potential analytic methods, 10 showed good correlation with NLR ($r^2 > 0.95$). By taking into account the degree of complexity of the methods, only 4 remain. The actual value of these methods needs to be evaluated in a large prospective response-monitoring study that compares results with patient outcome.

For response-monitoring studies in NSCLC, the nomograms in this study also provide the possibility of converting results obtained by one analytic method to another. This will allow comparison across studies (i.e., meta-analysis).

REFERENCES

1. Vansteenkiste JF, De Leyn PR, Deneffe GJ, et al. Survival and prognostic factors in resected N2 non-small cell lung cancer: a study of 140 cases—Leuven Lung Cancer Group. *Ann Thorac Surg.* 1997;63:1441–1450.
2. Misquitta A, Herman S, Winton T, Burkes R, Shepherd F, Rappaport D. Accuracy of CT in predicting prognosis in patients with non-small cell lung carcinoma after induction chemotherapy [abstract]. *Radiology.* 1994;193:257.
3. Romer W, Hanauske AR, Ziegler S, et al. Positron emission tomography in non-Hodgkin's lymphoma: assessment of chemotherapy with fluorodeoxyglucose. *Blood.* 1998;91:4464–4471.
4. Vansteenkiste JF, Stroobants SG, De Leyn PR, Dupont PJ, Verbeken EK. Potential use of FDG-PET scan after induction chemotherapy in surgically staged IIIa-N2 non-small-cell lung cancer: a prospective pilot study—The Leuven Lung Cancer Group. *Ann Oncol.* 1998;9:1193–1198.
5. MacManus MP, Hicks RJ, Wada M, et al. Early F-18 FDG-PET response to radical chemoradiotherapy correlates strongly with survival in unresectable non-small cell lung cancer [abstract]. *Proc Am Soc Clin Oncol.* 2000;19:483a.
6. Oshida M, Uno K, Suzuki M, et al. Predicting the prognoses of breast carcinoma patients with positron emission tomography using 2-deoxy-2-fluoro[^{18}F]-D-glucose. *Cancer.* 1998;82:2227–2234.
7. Hoekstra C, Paglianiti I, Hoekstra O, et al. Monitoring response to therapy in cancer using [^{18}F]-2-fluoro-2-deoxy-D-glucose and positron emission tomography: an overview of different analytical methods. *Eur J Nucl Med.* 2000;27:731–743.
8. Young H, Baum R, Cremerius U, et al. Measurement of clinical and sub-clinical tumour response using [^{18}F]-fluorodeoxyglucose and positron emission tomog-

raphy: review and 1999 EORTC recommendations. *Eur J Cancer*. 1999;35:1773–1782.
9. Boellaard R, van Balen SC, Lammertsma AA. Optimization of attenuation correction for PET studies of thorax and pelvis using count-based transmission scans [abstract]. *J Nucl Med.* 2001;42(suppl):196P.
10. Hoekstra C, Hoekstra O, Lammertsma AA. On the use of image derived input functions in oncological FDG PET studies. *Eur J Nucl Med.* 1999;26:1489–1492.
11. van der Weerdt AP, Klein LJ, Boellaard R, Visser CA, Visser FC, Lammertsma AA. Image-derived input functions for determination of MR_{glu} in cardiac ^{18}F-FDG PET scans. *J Nucl Med.* 2001;42:1622–1629.
12. Boellaard R, van Lingen A, Lammertsma AA. Experimental and clinical evaluation of iterative reconstruction (OSEM) in dynamic PET: quantitative characteristics and effects on kinetic modeling. *J Nucl Med.* 2001;42:808–817.
13. Patlak CS, Blasberg RG, Fenstermacher JD. Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data. *J Cereb Blood Flow Metab.* 1983;3:1–7.
14. Wu HM, Huang SC, Choi Y, Hoh CK, Hawkins RA. A modeling method to improve quantitation of fluorodeoxyglucose uptake in heterogeneous tumor tissue. *J Nucl Med.* 1995;36:297–306.
15. Zasadny KR, Wahl RL. Enhanced FDG-PET tumor imaging with correlation-coefficient filtered influx-constant images. *J Nucl Med.* 1996;37:371–374.
16. Wu HM, Hoh CK, Huang SC, Yao WJ, Phelps ME, Hawkins RA. Quantification of serial tumor glucose metabolism. *J Nucl Med.* 1996;37:506–513.
17. Hunter GJ, Hamberg LM, Alpert NM, Choi NC, Fischman AJ. Simplified measurement of deoxyglucose utilization rate. *J Nucl Med.* 1996;37:950–955.
18. Sadato N, Tsuchida T, Nakaumra S, et al. Non-invasive estimation of the net influx constant using the standardized uptake value for quantification of FDG uptake of tumours. *Eur J Nucl Med.* 1998;25:559–564.
19. Akaike H. A new look at the statistical identification. *IEEE Trans Automat Contr*. 1978;19:716–723.
20. Schwarz G. Estimating the dimension of a model. *Ann Statist.* 1978;6:461–464.
21. Shrout P, Fleiss J. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull.* 1979;86:420–428.
22. Schmidt K, Lucignani G, Moresco RM, et al. Errors introduced by tissue heterogeneity in estimation of local cerebral glucose utilization with current kinetic models of the [^{18}F]fluorodeoxyglucose method. *J Cereb Blood Flow Metab.* 1992;12:823–834.
23. Torizuka T, Tamaki N, Inokuma T, et al. In vivo assessment of glucose metabolism in hepatocellular carcinoma with FDG-PET. *J Nucl Med.* 1995;36:1811–1817.
24. Keyes JWJ. SUV: standard uptake or silly useless value? *J Nucl Med.* 1995;36:1836–1839.
25. 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.
26. Wahl RL, Henry CA, Ethier SP. Serum glucose: effects on tumor and normal tissue accumulation of 2-[F-18]-fluoro-2-deoxy-D-glucose in rodents with mammary carcinoma. *Radiology.* 1992;183:643–647.
27. 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.