RECIST AND BEYOND*

E. Coche[1]

The role of imaging in the evaluation of tumor response is expanding rapidly. The current response evaluation criteria in solid tumors (RECIST) based on anatomical changes suffers from many limitations related mainly to the inter- and intra-observer variability to delineate the tumoral edges. Consequently, there is a need to update and integrate the RECIST criteria beyond the classical anatomical changes with other more sophisticated methods using three-dimensional and functional criteria. The goal of this paper is to review the current criteria of RECIST measurements (RECIST 1.1) with their limitations and to evaluate the emerging solutions available with the new imaging techniques like PET-CT.

Key-word: Lung neoplasms, CT.

Quantification of tumor burden by medical imaging is being used with increasing frequency to assess the effectiveness of various anticancer therapies. Anatomic criteria defined as change in tumor size according to the World Health Organization (WHO) and the Response Evaluation Criteria in Solid Tumor (RECIST) criteria has long been considered as a surrogate marker of therapeutic efficacy. Recently other tumor parameters, beyond RECIST, including three-dimensional measurements, density changes, and avidity for FDG on PET-CT are considered as promising biomarkers to assess more rapidly the functional response to therapy.

The goal of this paper is to review the current criteria of RECIST measurements with their limitations and the emerging solutions available with the new imaging techniques.

Classical anatomical markers

RECIST criteria

Tumor response to therapy has been evaluated in many cancer clinical trials using two-dimensional anatomical criteria. In the late 1970s, the International Union against Cancer and the WHO introduced specific criteria for the codification of tumor response evaluation (1). Unfortunately, bidimensional measurements (i.e. the product of the longest diameter and its longest perpendicular diameter) to assess tumor burden response was found to have limited reproducibility (2). RECIST criteria were developed several years later (3) in order to provide an easier reproducible method of measurement with the concept that one-dimensional measurements were as informative as bidimensional measurements. Response to treatment was categorized into four main categories: complete response (CR), corresponding to a disappearance of all target lesions; partial response (PR), defined as a ≥ 30% decrease in tumor size from the baseline; progressive disease (PD), defined as a ≥ 20% increase in tumor size; and stable disease (SD), defined as small changes that do not meet the above criteria.

RECIST 1.1 criteria

New response evaluation criteria were published in 2009 (RECIST 1.1) (4) and include several changes compared to the previous version: the number of target lesions to assess tumour burden for response determination has been reduced from a maximum of ten to a maximum of five in total (and from five to two per organ, maximum). Lymph nodes with a short axis measuring ≥ 15 mm have been included as target lesions and included in the sum calculation of tumour response. New clarifications concerning disease progression has been made in addition to the previous definition of progression disease (20% increase in sum) for small lesions. New lesions documented by FDG-PET can be used as indicator of disease progression in the RECIST 1.1. The main differences between RECIST 1.0 and 1.1 and time point responses are summarized in table I and II respectively (4).

Limitations and difficulties of RECIST

There are many drawbacks with the RECIST criteria. When the tumoral lesion has variable morphology, uni-dimensional measurements may be inaccurate specifically when the lesion length exceeds twice its width (5).

Variability of lung tumor measurements represents also an important weakness of the RECIST method and may classify a patient in a wrong category due to those measurements errors. Oxnard et al. (6) determined the inter-measurement variance of CT for primary malignant lung lesions. Thirty patients with non-small cell lung carcinoma underwent non-enhanced CT, exited the scanner and were revaluated on the same scanner after a short delay. Images from both CT acquisitions were measured blindly by three radiologists. The radiologists manually measured the longest diameter of the target lesions on the two different scanners using a standard software. Lesions ranged from 1 to 8 cm in size. The absolute difference between scan measurements of single lesion ranged from 0 to 9 mm, with the greatest difference observed with the largest lesions and the greatest fractional differences observed with the smallest lesions. The potential impact of those measurements errors was simulated using statistical methods and found that aberrant assessments of partial response and progressive disease can occur as a result of measurement variance alone. For example, in this simulation, a 4-cm lesion has a measured range as a result of inter-measurement variance alone as broad as 3.5 to 4.5 cm, corresponding to approximately 12% change. Tumors with irregular edges, confluent or infiltrating boundaries pose the most significant challenges to data extraction and are highly observer dependent. Reliable diameter measurements in

---

Table I. — Comparison between RECIST 1.0 and RECIST 1.1.

	RECIST 1.0	RECIST 1.1
Minimum target lesion size	≥ 10 mm (Spiral CT) ≥ 20 mm (conventional CT, MRI)	≥ 10 mm (CT + MRI) ≥ 15 mm (lymph nodes)
N° of measurable lesions, max per organ	Max 10 5 per organ	Max 5 2 per organ
Measurement	Uni-dimensional Long axis for all lesions	Uni-dimensional Lymph-nodes = short axis
PD definition-Target	20% increase in SOD from Nadir	20% increase in SOD + min 5 mm increase from Nadir
PD definition-Non target	Unequivocal progression	Substantial worsening, tumor burden has increased sufficiently
Lymph node measurements	None	Measured short axis, ≥ 15 mm for target, ≥ 10 mm to < 15 mm for non-target, < 10 mm non-pathological
CR/PR confirmation	Required	Not required for randomized, controlled phase 3 trials

CR = Complete response
PR = Partial response
SD = Stable disease
PD = Progressive disease
NE = Non evaluable
SOD = sum of diameters
Nadir = The smallest sum on study

Modified from reference 4.

Table II. — Time point response: patients with target (± non-target) disease (ref. 4).

Target Lesions	Non-target lesions	New lesions	Overall response
CR	CR	No	CR
CR	Non-CR/Non-PD	No	PR
CR	Not evaluated	No	PR
PR	Non-PD or not all evaluated	No	PR
SD	Non-PD or not all evaluated	No	SD
Not all evaluated	Non-PD	No	NE
PD	Any	Yes or No	PD
Any	PD	Yes or No	PD
Any	Any	Yes	PD

CR = Complete response
PR = Partial response
SD = Stable disease
PD = Progressive disease
NE = Non evaluable.

"ground glass" opacities (Fig. 1), invasive lepidic carcinoma are especially problematic also, particularly if accompanied by pleural effusions.

Beyond RECIST

Three-dimensional evaluation

Recent advances in CT technology, specifically volumetric data acquisition and image processing, allows volumetric tumor burden quantification (7). Some preliminary studies have supported the use of three-dimensional measurements techniques for assessing tumor size (8). An important theoretical advantage of volumetric measurements is that simply estimating overall tumor in an organ can eliminate the limitation of measuring two target lesions per organ (RECIST criteria). In addition, volumetric measurement might be a better method to measure size changes of lesions that are confluent. Mozley and coworkers (9) have studied patients with lung cancers and have compared the reproducibility between long diameter and volume measurements. They obtained a lesser variability in volume measurement than one-dimensional measurements and conclude that measurements of change
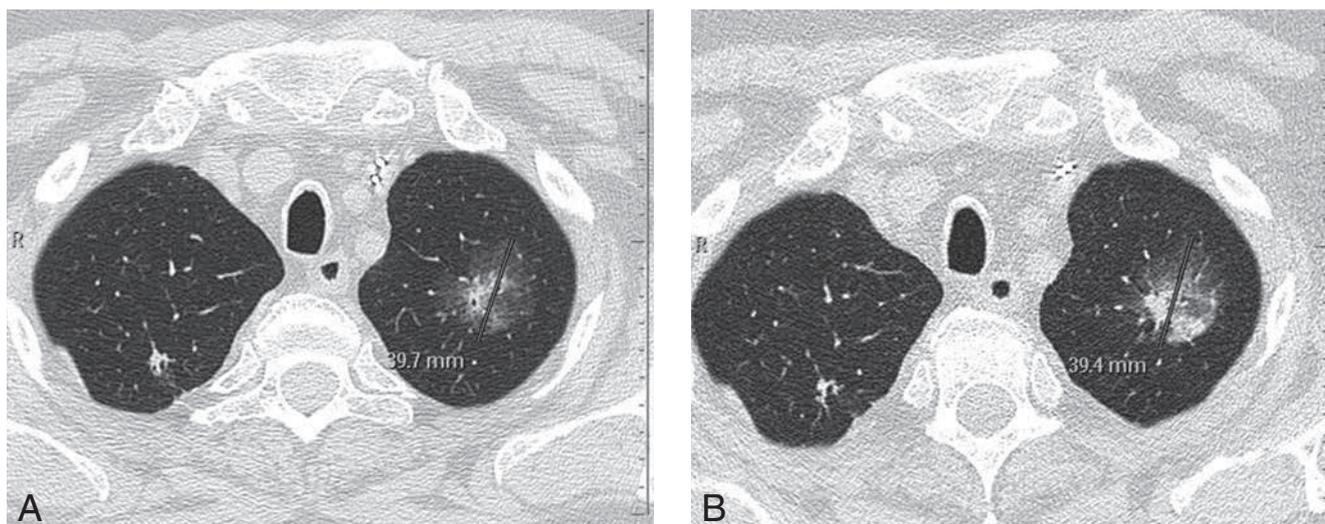
Fig. 1. — Limitations of RECIST criteria in semi-solid lesions.
82-year old man with a semi-solid lesion located at the left upper lobe.
A. Thin-slice CT performed at baseline revealed a solid lesion surrounded by "ground glass" opacity. Long axis of the whole lesion using RECIST criteria was measured at 40 mm. B. Thin-slice CT performed 6 months after baseline revealed a progression of the solid component. Long axis using RECIST criteria was unchanged. Clinical data were in favour of tumoral progression and the patient was operated. Pathology revealed a bronchial adenocarcinoma, classified pT1bN0.

in tumor volumes are adequately reproducible.

Density analysis

With the introduction of new cytostatic agents, central necrosis, density changes and cystic changes may occur before tumor shrinkage (Fig. 2). It has been suggested by several groups to include the measurement of density to the RECIST criteria on the basis of typical patterns of change observed in some categories of tumors and treatments. As an example, in gastrointestinal tumors, there is a decrease in tumor size at a lower magnitude and increase in tumor homogeneity and hypoattenuation with the treatment. A group from MD Alderson Cancer Center has suggested modifying the RECIST criteria by defining a 10% decrease in one-dimensional measurement or 15% decrease in density, as measured by Hounsfield units as a partial response (10).

PET-CT evaluation

The limitations of structural imaging modalities such as CT or MRI for accurately evaluating the response of tumors to non-surgical therapies are well known. Changes in tumor dimensions may occur slowly and incompletely. Biological parameters do change earlier, and these changes better reflect the actual tumor response. In this context, FDG-PET-CT imaging has a positive predictive value for N2 disease of 93%, compared to 66% for CT. The negative predictive value of PET is 75%, compared to 53% for CT (11). Moreover, a good metabolic response assessed by FDG-PET-CT is correlated with prolonged survival (12). Again, metabolic imaging is an exquisite method for the early quantitative assessment of the tumoral response (Fig. 3). As early as 2 days after the onset of treatment with gefitinib (an inhibitor of the EGF receptor), a decrease of FDG uptake can be measured by PET (13). This could help the clinician in deciding to discontinue a therapy in non-responding patients. Further studies are needed to better understand how FDG uptake reflects the multiple biological changes induced by these novel therapeutic agents.

Based on the literature supporting the use of 18F-FDG PET to assess early treatment response, quantitative PET criteria have been proposed to be used in clinical trials and possibly in clinical practice. Positron Emission Tomography Response Criteria in Solid Tumor (PERCIST) has been developed a few years ago and described extensively in a special issue of the Journal of Nuclear medicine in 2009 (14).

Multimodal integration

At present, many patients admitted for lung tumor work-up benefit from a multimodality approach combining a morphologic and functional imaging: MDCT, PET-CT, MR. The current challenge for the radiologist and the clinician resides in the integration of those different imaging modalities for an optimal exploitation of the data produced by the different sources. Many efforts are under way by several companies (Fig. 4: CT platform) to develop intelligent platform combining the different image modalities with fusion tools and different types of co-registration.
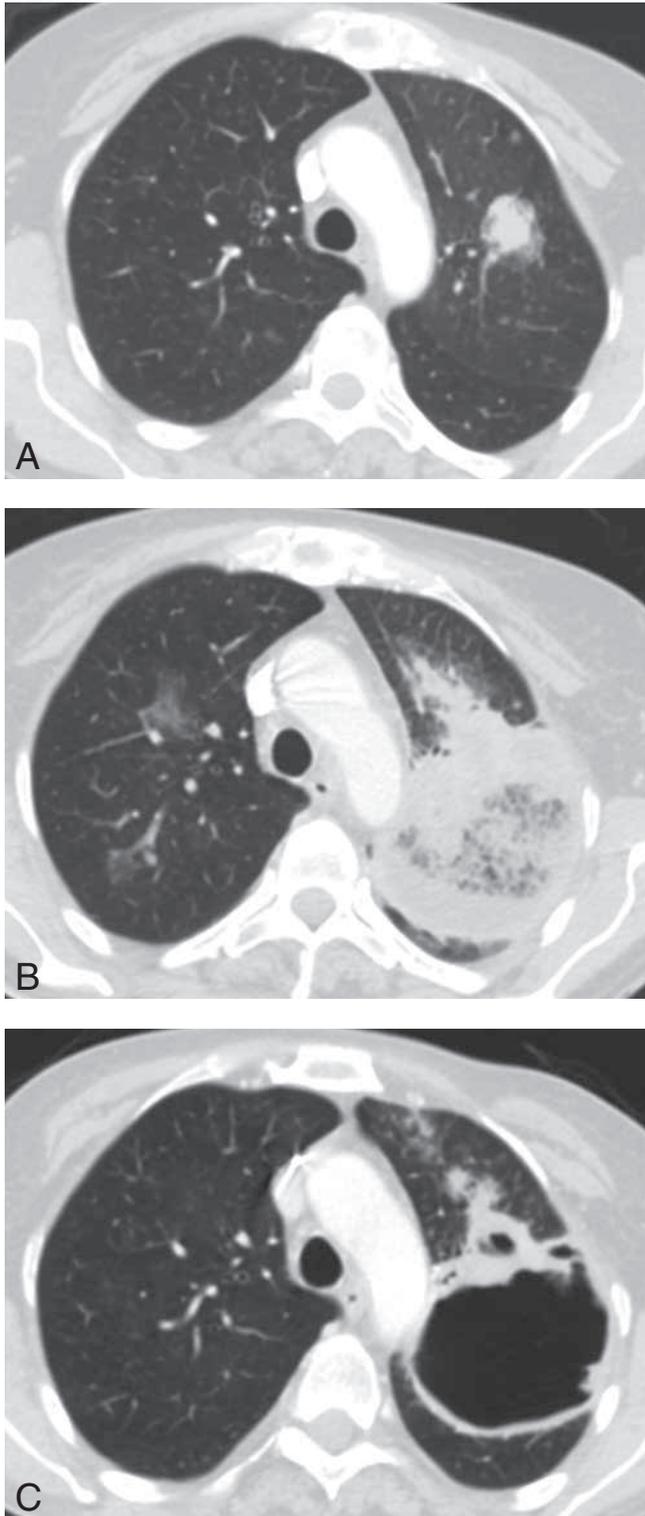
Fig. 2. — Inflammatory and cystic changes inappropriately assessed by RECIST criteria.

53-year old woman with bronchial adenocarcinoma of the left upper lobe. The patient was treated by chemotherapy and tyrosine kinase inhibitors (Sorafenib).

A. Spiral CT examination performed at baseline revealed an irregular lesion located at the left upper lobe. B. Spiral CT examination performed 1 month later revealed a significant increase of tumoral long axis probably related to inflammatory changes. C. Spiral CT examination performed 2 months later showed cystic changes. Morphological criteria following RECIST were judged as inappropriate to assess the response to therapy in this case.
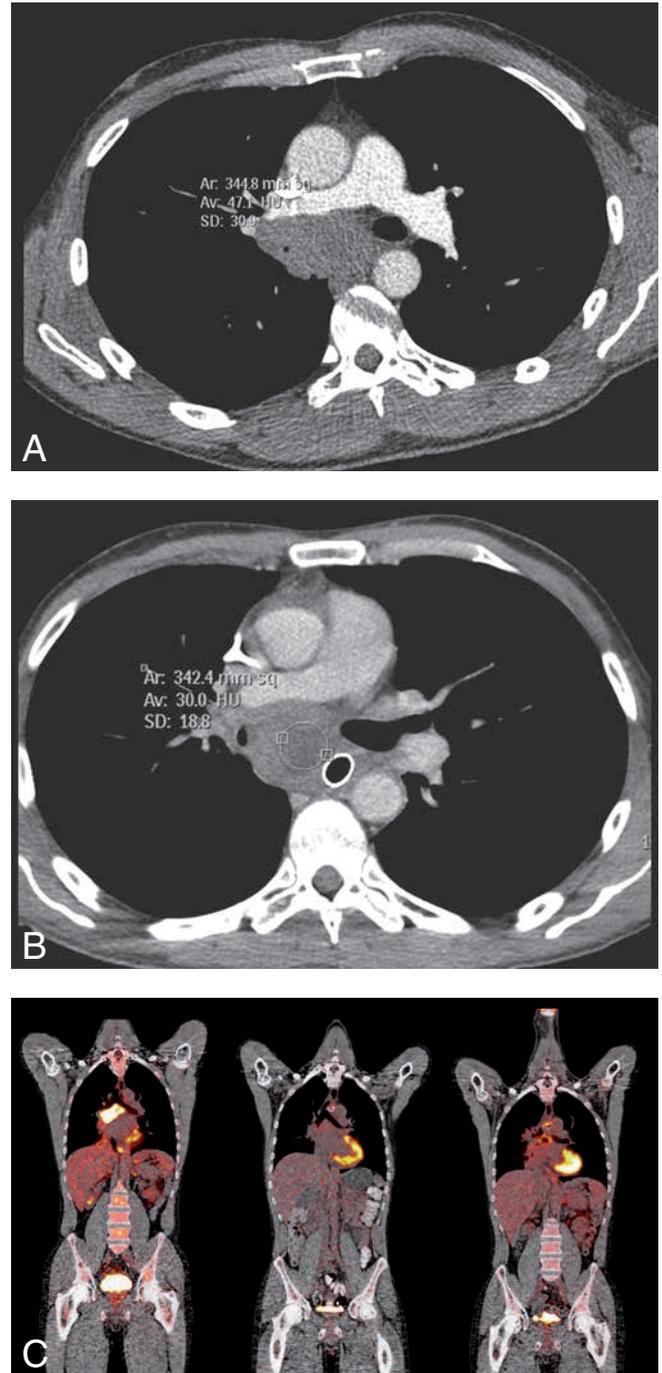
Fig. 3. — Current limitations of morphological markers (RE-CIST 1.1) to evaluate the response of a tumour to therapy.

A 46-year old man with NSCLC (squamous cell tumor) of the right inferior lobe initially staged as cT3N2M0 and treated with chemotherapy (Cisplatine, VP16) and radiotherapy.

A. MDCT performed after intravenous contrast medium injection. Axial slice obtained at the level of the subcarina area (December 2010). A large subcarinal (station 7) is observed. Its density is homogeneous. B. MDCT performed after intravenous contrast medium injection. Axial slice obtained at the level of the subcarina area (March 2011). Note an oesophageal stent in relationship with post-radiotherapy esophagitis with severe stenosis. The subcarinal lymphadenopathy has decreased in density. Its short axis has slightly decreased but non significantly following RECIST 1.1. C. 18FDG- PET-CT acquired on December 2010, April 2011 and august 2011 respectively, showing the functional response to the therapy. On april 2011, The subcarinal lymph node has clearly reduced is SUV.

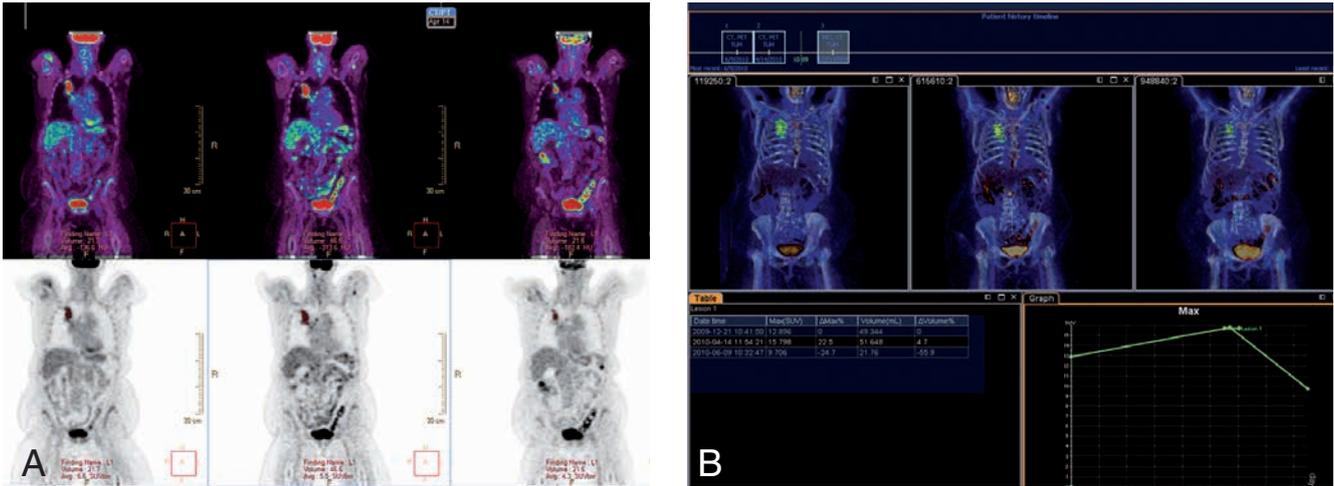Courtesy of Dr F.-X. Hanin, Nuclear Medicine Unit, Cliniques Universitaires St-Luc, Brussels, Belgium.

Fig. 4. — Integration of multimodality tumoral response.

The Multimodality Tumor Tracking application (Philips healthcare, Cleveland, OH, US) offers tools to assist clinicians in monitoring change in disease status including disease progression or assessment of therapy response using sequential PET/CT, SPECT/CT, MR, and CT exams, with automatic segmentation of target lesions and quantified results over time.

A. Automatic delineation of the right upper lung tumor (in red) on PET-CT. Display of the SUV below each image at every time point. B. The modality provides an integration of PET-CT and volumetric CT with a table summarizing the different tumor volumes, SUV values at different time points. A schematic representation of the tumoral measurements is provided at the left lower corner of the screen.

References

1. World Health Organization. Who Handbook for Reporting Results of Cancer Treatment. Offset Publication, Geneva, Switzerland, 1979.
2. Park J.O., Lee S.I., Song S.Y., et al.: Measuring response in solid tumors: comparison of RECIST and WHO response criteria. *Jpn J Clin Oncol,* 2003, 33: 533-537.
3. Therasse P., Arbuck S.G., Eisenhauer E.A., et al.: New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J Nat Cancer Inst*, 2000, 92: 205-216.
4. Eisenhauer E.A., Therasse P., Bogaerts J., et al.: New response evaluation criteria in solid tumors: revised RECIST guidelines (version 1.1). *Eur J Cancer*, 2009, 45: 228-247.
5. Spears C.P.: Volume doubling measurement of spherical and ellipsoïdal tumors. *Med Pediatr Oncol*, 1984, 12: 212-217.
6. Oxnard G.R., Zhao B., Sima C., et al.: Variability of lung tumor measurements on repeat computed tomography scans taken within 15 minutes. *J Clin Oncol*, 2011, 29: 3114-3119.
7. Zeman R.X., Fox S.H., Silverman P.M., et al.: Helical (spiral) CT of the abdomen. *Am J Roentgenol,* 1993, 160: 719-725.
8. Hopper K.D., Kasales C.J., Eggli K.D., et al.: The impact of 2D versus 3D quantification of tumor bulk determination on current methods of assessing response to treatment. *J Comput Assist Tomogr*, 1996, 20: 930-937.
9. Mozley P.D., Bendtsen C., Zhao B., et al.: Measurement of tumor volumes improves RECIST-based response assessment in Advanced lung cancer. *Translational Oncology*, 2012, 5: 19-25.
10. Choi H.: Critical issues in response evaluation on computed tomography: lessons from gastrointestinal stromal tumor model. *Curr Oncol Rep*, 2005, 7: 307-311.
11. De Leyn P., Stroobants S., De Wever W., et al.: Prospective comparative study of integrated positron emission tomography-computed tomography scan compared compared with remediastinoscopy in the assessment of residual mediastinal lymph node disease after induction chemotherapy for mediastinoscopy-proven stage IIIA-N2 Non-small-cell lung cancer: a Leuven Lung Cancer Group Study. *J Clin Oncol,* 2006, 24: 3333-3339.
12. Eschmann S.M., Friedel G., Paulsen F., et al.: Repeat 18F-FDG PET for monitoring neoadjuvant chemotherapy in patients with stage III non-small cell lung cancer. *Lung Cancer,* 2007, 55: 165-171.
13. Sunaga N., Oriuchi K., Kaira K., et al.: Usefulness of FDG-PET for early prediction of the response to gefitinib in non-small cell lung cancer. *Lung Cancer,* 2008, 59: 203-210.
14. Wahl R.L., Jacene H., Kasamon Y., Lodge M.A.: From RECIST to PERCIST: Evolving considerations for PET response criteria in solid tumors. *J Nucl Med,* 2009, 50: 122S-150S.