# Tumor response to chemotherapy: The validity and reproducibility of RECIST guidelines in NSCLC patients[1]

**Hirokazu Watanabe,[2, 3, 5] Seiichiro Yamamoto,[4] Hideo Kunitoh,[2] Ikuo Sekine,[2] Noboru Yamamoto,[2] Yuichiro Ohe,[2] Tomohide Tamura,[2] Tetsuro Kodama,[2] Kazuro Sugimura[3] and Nagahiro Saijo[2]**

[2]Department of Medical Oncology, National Cancer Center Hospital and [4]Cancer Information and Epidemiology Division, National Cancer Center Research, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045; and [3]Department of Radiology, Kobe University Graduate School of Medicine, 7-5-2 Kusunoki-cho, Chuo-ku, Kobe 650-0017

(Received July 10, 2003/Revised September 25, 2003/Accepted October 6, 2003)

**We investigated the validity and inter-criteria reproducibility between RECIST (Response Evaluation Criteria in Solid Tumors) guidelines and WHO (World Health Organization) criteria, considering the decrease in patient numbers resulting from inclusion of the minimum lesion size criterion introduced in RECIST guidelines. RECIST guidelines are based on unidimensional measurement and exclusion of small lesions from measurement. The aims of the study were to examine: (1) the effect of the minimum lesion size criterion, (2) the validity of unidimensional and bidimensional measurements, i.e., their relationship with tumor volume, (3) the inter-criteria reproducibility between current RECIST guidelines and previous WHO criteria. One hundred and twenty patients with non-small cell lung cancer (NSCLC) in clinical trials were evaluated. By applying the minimum lesion size criterion, six cases became ineligible without any influence on precision of tumor volume measurement. In the validity study, actual tumor volume was regarded as the gold standard. Although the unidimensional measurement had a lower correlation with tumor volume value than the bidimensional measurement, both the unidimensional measurement and bidimensional measurement correlated sufficiently well with tumor volume changes and the assessed tumor volume response. In the inter-criteria reproducibility study between RECIST guidelines and WHO criteria, the response rate assessed by RECIST guidelines (19.3%) was almost the same as that assessed by WHO criteria (20.0%). In conclusion, RECIST guidelines are adequate for evaluating tumor response to chemotherapy in terms of both validity in relation to tumor volume and inter-criteria reproducibility with the WHO criteria. (Cancer Sci 2003; 94: 1015–1020)**

New guidelines for evaluating tumor response, RECIST (Response Evaluation Criteria in Solid Tumors) guidelines, have been recently adopted by many organizations.[1)] RECIST guidelines stipulate the use of unidimensional measurement of lesions in contrast with the bidimensional measurement stipulated by WHO (World Health Organization) criteria[2)] and define the minimum lesion size allowable for measurability of the lesion to be no less than double the slice thickness on computed tomography (CT) or magnetic resonance imaging (MRI). When this minimum lesion size is included in the eligibility criteria, the number of patients with measurable lesions decreases in comparison to previous WHO criteria, because some patients with only small lesions are excluded from the eligibility criteria.

Several previous studies have demonstrated the inter-measurement reproducibility between unidimensional and bidimensional measurement in the same cases.[1, 3–5)] However, they have not considered the decrease in number of eligible cases as a result of the inclusion of the minimum lesion size criterion, and thus they have been unable to demonstrate true inter-criteria reproducibility between RECIST guidelines and WHO criteria. In addition, validity has been based on the subjective theoretical inference that unidimensional measurement is more proportional to the logarithm of cell numbers than bidimensional measurement, but this hypothesis has not been objectively evaluated.[6)]

Before introducing RECIST guidelines at our institution, we considered that the validity and inter-criteria reproducibility between the new and conventional criteria should be investigated. We had three objectives in investigating whether RECIST guidelines were adequate for evaluating tumor response to chemotherapy. These were to assess:

(1) The effect of the minimum lesion size criterion on the number of eligible patients and on the precision of tumor volume measurement.

(2) The validity of RECIST guidelines and WHO criteria by correlating the two different dimensional measurements with tumor volume as the gold standard, i.e. by correlating the relationship with tumor volume, and by applying the minimum lesion size criterion to these measurements.

(3) The inter-criteria reproducibility between current RECIST guidelines (unidimensional measurement in measurable cases excluding small lesions) and previous WHO criteria (bidimensional measurement in all cases including small lesions).

## Materials and Methods

**Patient population.** This is a retrospective study of radiological findings of patients who underwent chemotherapy in clinical trials for advanced non-small cell lung cancer (NSCLC). The subjects were patients treated at the Medical Oncology Division of the National Cancer Center Hospital in Tokyo, between January 1996 and April 2000. All clinical trials were conducted according to the Helsinki Declaration and the protocol was approved by the local ethics committee. Written informed consent was obtained from each patient for each treatment protocol, which included the secondary use of treatment-associated documents. Patients were staged according to the UICC TNM Classification of malignant tumors.[7)]

One hundred and twenty patients in clinical trials who fulfilled the following criteria were selected for the study:

[1]A part of this work was presented at the 36th ASCO Annual Meeting, May 23, 2000.

[5]To whom correspondence should be addressed at Department of Radiology, Kobe University Graduate School of Medicine.
E-mail: h-watanabe@kcc.zaq.ne.jp

Abbreviations: WHO, World Health Organization; RECIST, Response Evaluation Criteria in Solid Tumors; NSCLC, non-small cell lung cancer; CDDP, cisplatin; CT, computed tomography; MRI, magnetic resonance imaging; CR, complete response; PR, partial response; SD, stable disease; PD, progressive disease; CI, confidence intervals.

**Table 1. Patient and characteristics (all 120 cases and 50 cases analyzed by area volume)**

| | | All | Analyzed by area volume |
|---|---|---|---|
| No. of patients | | 120 | 50 |
| Age | Median | 60 | 62 |
| | Range | 38–75 | 44–75 |
| Sex | Male | 82 | 35 |
| | Female | 38 | 15 |
| Stage | IIIB | 41 | 11 |
| | IV | 79 | 39 |
| Histology | Adeno ca. | 88 | 39 |
| | Squamous cell ca. | 21 | 8 |
| | Adenosquamous ca. | 1 | 1 |
| | Large-cell ca. | 10 | 2 |
| Regimen | Cisplatin and paclitaxel | 35 | 14 |
| | Cisplatin and vindesine | 26 | 13 |
| | Cisplatin, docetaxel and ifosfamide | 21 | 8 |
| | Cisplatin and irrinotecan | 12 | 0 |
| | Cisplatin and docetaxel | 13 | 9 |
| | Cisplatin, navelbine and mitomycin | 4 | 0 |
| | Cisplatin, vindesine and mitomycin | 3 | 0 |
| | Cisplatin and gemcitabine | 4 | 4 |
| | Cisplatin and navelbine | 2 | 2 |

1. They were histologically or cytologically diagnosed with NSCLC.

2. They were treated with cisplatin (CDDP)-based chemotherapy in clinical trials.

3. They had at least one measurable lesion.

4. They had undergone CT scans periodically for evaluating tumor response to chemotherapy prior to and at least once after treatment.

The patients' characteristics were as follows: male/female=82/38, median age=60 (range 38–75), stage III B/IV=41/79. Chemotherapy regimens are listed in Table 1.

Patients treated in daily clinical practice were considered to be unsuitable and were excluded from this study, as tumor response evaluation in the daily clinical practice of oncology is not always performed according to predefined criteria, but rather is made by subjective medical judgment based on clinical and laboratory data. In addition, tumor response evaluation is not always performed on the basis of CT examinations, and the intervals between tumor evaluations can be irregular.

**Image analysis.** Almost all images were acquired with a TCT-900S Superhelix (Toshiba Medical, Tokyo), with the remainder having been scanned on an X-Vigor helical CT scanner (Toshiba Medical). Helical CT was performed with fixed scanning parameters including 120 kVp, 200 mAs, table speed of 15 mm/sec (pitch, 1.5:1), 1 second per rotation and contrast agent throughout baseline and follow-up evaluations. Image reconstruction was performed at intervals of 10 mm.

We selected the unidimensional value as the longest diameter of a tumor, the bidimensional value as the product of the unidimensional value and the longest diameter perpendicular to it, the tridimensional volume value as the product of the bidimensional value and tumor height, and the area volume value as the integration of tumor area. In addition, unidimensional change, bidimensional change, tridimensional volume change, and area volume change were calculated as percentage changes in tumor size from the baseline evaluation to the follow-up evaluation. Three hundred and fifty-two evaluations were performed in 120 cases, which included 120 baseline and 232 follow-up evaluations.

Two types of CT-assisted tumor volume measurement believed to give values very close to the true tumor volume were employed and calculations were based on digitized images measured using electronic calipers (Fig.1 and Table 2). First, tridimensional volume was calculated as the product of the unidimensional value, the longest diameter perpendicular to it and tumor height. Second, area volume measurement was performed. The tumor area was measured by manually tracing the tumor outline with a computer mouse on each axial slice in which the tumor was visualized, and multiplying it by the slice thickness to yield a slice volume. The individual slice volumes were then added together to obtain an overall volume.[8)]

In all 120 cases, we measured three parameters (the unidimensional value, the largest diameter perpendicular to it and the tumor height). A new computer system, which could measure tumor area on a terminal monitor, was introduced at our institution in 1999. Thus for 50 cases entering from January

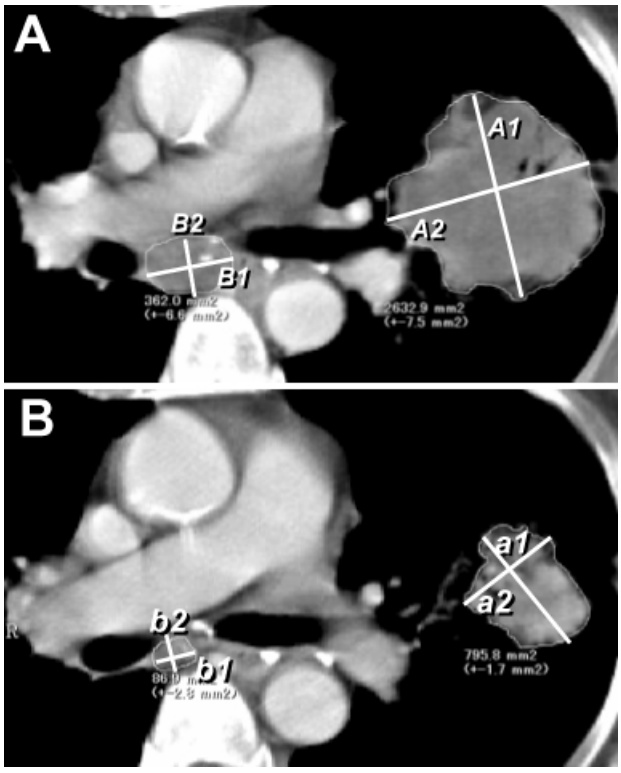


**Fig. 1.** Tumor measurement method: baseline evaluation (A) and follow-up evaluation (B).

1999, we were able to measure four parameters (the unidimensional value, the largest diameter perpendicular to it, the tumor height and the tumor area) in each lesion, and, as a result, the tumor volume could be more accurately calculated in these cases.

Tumors were retrospectively measured at baseline evaluation (obtained before the initiation of chemotherapy) and at regular intervals during the trials. All baseline and follow-up evaluations were retrospectively measured by the same radiologist (H.W.), who was blinded to the patient files. Lung lesions and mediastinal lesions were estimated on CT images mainly using soft tissue windows. Metastatic lesions of the abdomen and the brain were also assessed by CT examinations. If there were two or more lesions, the sum of all lesions (primary lesion, mediastinal and hilar lymphadenopathy, and metastasis lesions) up to a maximum of 5 lesions per organ and 10 lesions in total was calculated.

**Tumor response evaluation.** Tumor response evaluation was categorized into complete response (CR), partial response (PR), stable disease (SD) and progressive disease (PD) based on RECIST guidelines and WHO criteria. The RECIST PR is defined as a 30% decrease in the sum of the longest diameters and the WHO PR is defined as a 50% decrease in the sum of the products. The RECIST PD is defined as a 20% increase in the sum of the longest diameters and the WHO PD is defined as a 25% increase in the sum of the products of all lesions or in the product of any one lesion.

Each patient's tumor measurements were also evaluated according to volume criteria. If we consider spherical tumors, which grow and shrink isometrically, both the WHO 50% decrease and the RECIST 30% decrease result in a 65% tumor volume decrease. In addition, the volume change required to qualify for PD is not equivalent in the two sets of criteria. The RECIST PD (a 73% increase in tumor volume) requires a larger tumor increase than the WHO PD (a 40% increase in tumor volume).

In both criteria, a minimum interval of 4 weeks is required to confirm CR or PR. In the case of SD in RECIST guidelines, measurements must meet the SD criteria at least once after study entry at a minimum interval. In the present study, for correlation with WHO criteria, this minimum interval criterion was not applied.

**The effect of minimum lesion size criteria.** We examined the impact on the number of eligible patients by the minimum lesion size criterion introduced in RECIST guidelines, requiring a lesion whose minimum size is no less than double the slice thickness on images. The slice thickness was 10 mm in the present study, so the minimum lesion size was required to be no less than 20 mm at baseline evaluation before treatment and small lesions were defined as lesions less than 20 mm. The measurable cases were defined to exclude cases with only small lesions from all cases. We defined RECIST guidelines in terms of unidimensional measurement in measurable cases excluding small lesions and WHO criteria in terms of bidimensional measurement in all cases, including small lesions.

The impact of the minimum lesion size criterion on the precision of tumor volume measurements was evaluated by comparing the standard error of the correlation coefficient between measurable cases (excluding small lesions) and all cases (including small lesions).

**The validity of unidimensional and bidimensional measurements, i.e. the relationship with tumor volume.** To examine the validity of RECIST guidelines and WHO criteria, we estimated the Spearman's correlation coefficients between the two different dimen-

**Table 2. Tumor measurement method shown in Fig. 1**

| | Value | Change |
|---|---|---|
| Unidimensional measurement | A1+B1+... | a1+b1+.../A1+B1+... |
| Bidimensional measurement | A1×A2+B1×B2+... | a1×a2+b1×b2+.../A1×A2+B1×B2+... |
| Tridimensional volume | A1×A2×A3*+B1×B2×B3+... | a1×a2×a3+b1×b2×b3+.../A1×A2×A3*+B1×B2×B3+... |
| Area volume | (Area A+Area B+...)×Slice thickness | Volume a+Volume b+.../Volume A+Volume B+... |

* A3: tumor height.

**Table 3. Validity (Tridimensional volume measurement)**

| | Value | | | Change | | | Response rate | | |
|---|---|---|---|---|---|---|---|---|---|
| | Uni | Bi | Evaluations | Uni | Bi | Evaluations | Uni | Bi | Volume |
| Measurable 114 cases | 0.85 (0.018)* | 0.97 (0.005)* | 336 | 0.9 (0.02)* | 0.95 (0.010)* | 222 | 19.30% (12.5–27.7)** | 21.10% (14.0–29.7)** | 19.30% (12.5–27.7)** |
| All 120 cases | 0.84 (0.019)* | 0.97 (0.005)* | 352 | 0.9 (0.016)* | 0.94 (0.011)* | 232 | 19.20% (12.6–27.4)** | 20.00% (13.3–28.3)** | 19.20% (12.6–27.4)** |

* Spearman's correlation coefficient (standard error).
** The response rate (95% confidence intervals).
Abbreviations: Uni represents the unidimensional measurement; Bi represents the bidimensional measurement.

**Table 4. Validity (Area volume measurement)**

| | Value | | | Change | | | Response rate | | |
|---|---|---|---|---|---|---|---|---|---|
| | Uni | Bi | Evaluations | Uni | Bi | Evaluations | Uni | Bi | Volume |
| Measurable 46 cases | 0.8 (0.040)* | 0.92 (0.019)* | 133 | 0.84 (0.036)* | 0.85 (0.038)* | 87 | 19.60% (9.4–39.9)** | 19.60% (9.4–39.9)** | 17.40% (7.8–31.4)** |
| All 50 cases | 0.8 (0.038)* | 0.93 (0.017)* | 144 | 0.83 (0.032)* | 0.81 (0.044)* | 94 | 18.00% (8.5–31.4)** | 18.00% (8.5–31.4)** | 16.00% (7.2–29.1)** |

* Spearman's correlation coefficient (standard error).
** The response rate (95% confidence intervals).
Abbreviations: Uni represents the unidimensional measurement; Bi represents the bidimensional measurement.

sional values (unidimensional measurement and bidimensional measurement) and the gold standard value (tridimensional volume measurement and area volume measurement). We also estimated the Spearman's correlation coefficient between the two different dimensional changes and the gold standard changes. Furthermore, we compared tumor responses assessed by using the two different dimensional criteria with those using the gold standard criteria.

**The inter-criteria reproducibility between RECIST guidelines and WHO criteria.** To examine the inter-criteria reproducibility between RECIST guidelines and WHO criteria, the tumor responses assessed by applying the two criteria were divided into four categories which represented CR, PR, SD and PD. We examined whether the response rate would change as a result of unidimensional measurement or application of the minimum lesion size criterion.

All analyses were conducted with SAS ver.8.02 (SAS Institute, Cary, NC).

## Results

**The effect of the minimum lesion size criterion.** When the minimum lesion size criterion for measurable lesions introduced in RECIST guidelines was applied, six cases (5%) out of 120 cases turned out to have no measurable lesions and were considered ineligible for tumor response evaluation. The number of eligible cases thus decreased from 120 to 114. Additionally, in 40 of these 114 cases, the number of measurable lesions decreased. There was no influence on the number of measurable lesions in 74 cases.

We also examined the effect on the precision of tumor volume measurements when the minimum lesion size criterion in RECIST guidelines was applied and found that the standard error of the correlation coefficient between measurable cases (excluding small lesions) and all cases (including small lesions) was almost the same (Tables 3 and 4).

**The validity of unidimensional and bidimensional measurements, i.e. the relationship with tumor volume.** In the validity examinations, the two types of CT-assisted tumor volume measurement (tridimensional volume measurement and area volume measurement) were regarded as the gold standard. Table 3 shows Spearman's correlation coefficient and the standard error for each evaluation. The response rate for each evaluation is also shown in Table 3.

The unidimensional value had a lower correlation with the tridimensional volume value than the bidimensional value (Figs. 2 and 3). However, as regards the correlation with the tridimensional volume change, unidimensional change exhibited no difference from bidimensional change. In measurable cases excluding small lesions, the response rates were 19.3% (95% CI (confidence intervals): 12.5–27.7%) (22/114) for unidimensional measurement (RECIST guidelines) and 19.3% (95% CI: 12.5–27.7%) (22/114) for tridimensional volume measurement. In all cases including small lesions, the response rates were 20.0% (95% CI: 13.3–28.3%) (24/120) for bidimensional measurement (WHO criteria) and 19.2% (95% CI: 12.6–27.4%) (23/120) for tridimensional volume measurement. The response rates among unidimensional measurement, bidimensional measurement and tridimensional volume measurement were almost the same.

Area volume measurement showed the same tendency as the tridimensional volume measurement (Table 4). Unidimensional value had a lower correlation with the area volume value than bidimensional value. In terms of the correlation with area vol-

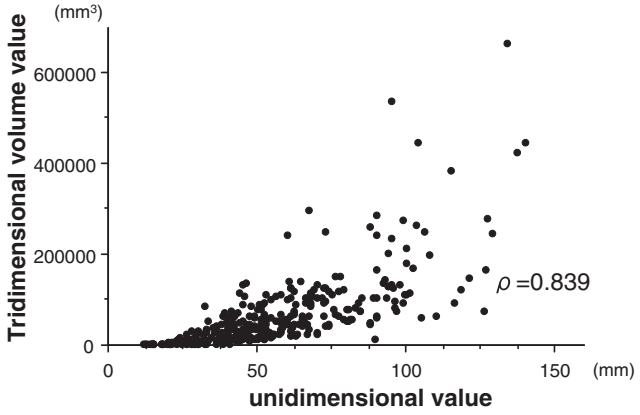


**Fig. 2.** Validity: Correlation between unidimensional value and tridimensional volume value in all 120 cases (352 evaluations).

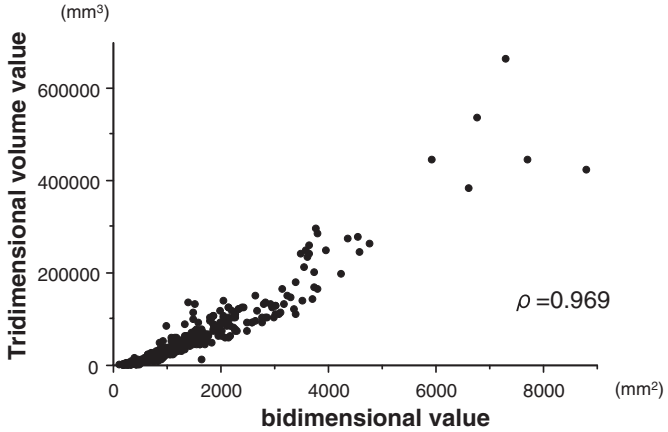


**Fig. 3.** Validity: Correlation between bidimensional value and tridimensional volume value in all 120 cases (352 evaluations).

**Table 5. Inter-criteria reproducibility (Comparison of tumor response evaluations)**

A. Unidimensional measurement

| | CR | PR | SD | PD | Ineligible | Response rate |
|---|---|---|---|---|---|---|
| Measurable cases (RECIST) | 0 | 22 | 77 | 15 | 6 | 19.3% (22/114) |
| All cases | 0 | 23 | 82 | 15 | 0 | 19.2% (23/120) |

B. Bidimensional measurement

| | CR | PR | SD | PD | Ineligible | Response rate |
|---|---|---|---|---|---|---|
| Measurable cases | 0 | 24 | 71 | 19 | 6 | 21.1% (24/114) |
| All cases (WHO) | 0 | 24 | 75 | 21 | 0 | 20.0% (24/120) |

Abbreviations: CR, complete response; PR, partial response; SD, stable disease; PD, progressive disease.

ume change, however, unidimensional change showed no difference from bidimensional change. In measurable cases excluding small lesions, the response rates were 19.6% (95% CI: 9.4–33.9%) (9/46) for unidimensional measurement (RECIST guidelines) and 17.4% (95% CI: 7.8–31.4%) (8/46) for area volume measurement. In all cases including small lesions, the response rates were 18.0% (95% CI: 8.5–31.4%) (9/50) for bidimensional measurement (WHO criteria) and 16.0% (95% CI: 7.2–29.1%) (8/50) for area volume measurement. Thus, the response rates among unidimensional measurement, bidimensional measurement and area volume measurement were almost the same.

**The inter-criteria reproducibility between RECIST guidelines and WHO criteria.** In the inter-criteria reproducibility examination, the response rates based on RECIST guidelines and WHO criteria were evaluated (Table 5). In the unidimensional measurement, the best response (measurable cases/all cases) was CR 0/0, PR 22/23, SD 77/82, and PD 15/15, while the response rates were 19.3% (95% CI: 12.5–27.7%) (22/114) for measurable cases and 19.2% (95% CI: 12.5–27.4%) (23/120) for all cases. In the bidimensional measurement, the best response (measurable cases/all cases) was CR 0/0, PR 24/24, SD 71/75, and PD 19/21, while the response rates were 21.1% (95% CI: 14.0–29.7%) (24/114) for measurable cases and 20.0% (95% CI: 13.3–28.3%) (24/120) for all cases. Thus, the response rate was almost the same for RECIST guidelines and WHO criteria at 19.3% and 20.0%, respectively.

Unidimensional value and bidimensional value also correlated well (all cases 0.91/measurable cases 0.92), as did unidimensional change and bidimensional change (all cases 0.93/measurable cases 0.93).

Eleven patients developed new lesions. Four cases were assessed as PD by RECIST guidelines but six cases were assessed as PD by the WHO criteria due to an increase in the sum of all pre-existing lesions. In addition, four cases were assessed as PD by the WHO criteria due to an increase in any one pre-existing lesion. Thus, in total, 15 cases were assessed as PD by RECIST guidelines (unidimensional measurement in measurable cases) and 21 cases by the WHO criteria (bidimensional measurement in all cases).

## Discussion

To our knowledge, this is the first statistical analysis of actual measurements that can clearly show the validity and inter-criteria reproducibility between RECIST guidelines and WHO criteria, considering the decrease in patient numbers resulting from inclusion of the minimum lesion size criterion.

When the minimum lesion size criterion was applied, the eligible cases changed from 120 to 114 (95%) cases. Thus, we had to try to recruit 127 cases in total to evaluate 120 eligible cases, i.e., 7 cases (5.8%) more than previously needed had to be recruited.

The minimum lesion size criterion, i.e., evaluation of only measurable lesions excluding small lesions, could not be considered to have any influence on the precision of tumor volume measurement. However, the role of the minimum lesion size may require further examination (for example, by considering inter-observer reproducibility).

In the validity study, actual tumor volume was regarded as the gold standard. Although the unidimensional measurement had a lower correlation with tumor volume value than the bidimensional measurement, both the unidimensional measurement and bidimensional measurement correlated sufficiently well with tumor volume changes and the assessed tumor volume response. These results led to the conclusion that both RECIST guidelines and WHO criteria were valid in relation to tumor volume.

In the inter-criteria reproducibility study between RECIST guidelines and WHO criteria, the response rate assessed by applying RECIST guidelines (19.3%) was almost the same as that assessed by applying WHO criteria (20.0%). These results led to the conclusion that there was sufficient inter-criteria reproducibility between RECIST guidelines and WHO criteria.

The bidimensional value offers advantages in assessment of tumor volume over the unidimensional value because it has the potential to provide a more accurate description of tumor volume. However, the differences resulting from the measurement method are not large enough to alter assessed tumor change or to affect the categorization of tumor response. In terms of tumor response classification into only four categories (CR, PR, SD, PD), the response rates obtained by all measurements showed a high agreement. This is why RECIST guidelines using only simple unidimensional measurement are adequate for accurately evaluating tumor response to chemotherapy.

In this study, there were more WHO PD cases (21 cases) than RECIST PD cases (15 cases) because of the differences in definition of PD in the two sets of criteria. Two cases were assessed as PD by the WHO criteria, but not by RECIST guidelines due to increases in all lesions (a 40% increase vs. a 73% increase in tumor volume). In addition, four cases were assessed as PD by the WHO criteria due to increases in any one lesion, although other target lesions had not progressed and tumor response evaluation could be assessed SD by RECIST guidelines based on the sum of the products of all lesions. This was a common problem with the WHO criteria.

CT-assisted tumor volume calculations with helical CT are likely to provide the most precise measurements when assessing irregularly shaped structures which exhibit non-uniform size changes. Simple tridimensional volume measurements can be as accurate as area volume measurements.[9, 10] This was reflected in our study, in which tridimensional volume measurements and area volume measurements were highly correlated (all cases 0.98/measurable cases 0.97).

However, the present study had several weaknesses and some further investigation is needed. First, nontarget lesions could not be accurately evaluated in this study because it is retrospective. However, as there was no CR in the target lesions, the best overall response was not influenced and the same conclusions could be reached without the evaluation of nontarget lesions. Second, the observer was a single radiologist (H.W.), so the influence of inter-observer reproducibility and intra-observer reproducibility could not be examined. Third, tumor volume was regarded as the gold standard and as a surrogate for survival. As the true gold standard is survival, further studies to correlate tumor response with survival are needed in larger trials. Fourth, our slice thickness was 10 mm, and therefore 20 mm were defined as minimum lesion size. However, RECIST guidelines allow for a minimum lesion size of 10 mm if measurements are made with a slice thickness of 5 mm with helical CT. The multidetector-row CT system, which can create a thinner slice thickness, is a recent development in routine clinical practice. The outcomes for currently ineligible patients when applying the thinner slice thickness should be evaluated in a further study.

We conclude that RECIST guidelines are adequate for evaluating tumor response to chemotherapy, both in relation to tumor volume and inter-criteria reproducibility with the WHO criteria. Thus, the present study serves to support and strengthen the simplification and standardization of tumor response evaluation to chemotherapy offered by the RECIST guidelines.

The authors grateful thank Patrick Therasse (Director of EORTC Data Center) for kindly reviewing the manuscript and Masahiko Kusumoto (Department of Radiology, National Cancer Center Hospital) for comments on radiology.

1. Therasse P, Arbuck SG, Eisenhauer EA, Wanders J, Kaplan RS, Rubinstein L, Verweij J, Van Glabbeke M, van Oosterom AT, Christian MC, Gwyther SG. New guidelines to evaluate the response to treatment in solid tumors. *J Natl Cancer Inst* 2000; **92**: 205–16.
2. Miller AB, Hoogstraten B, Staquet M, Winkler A. Reporting results of cancer treatment. *Cancer* 1981; **47**: 207–14.
3. Sohaib SA, Turner B, Hanson JA, Farquharson M, Oliver RT, Reznek RH. CT assessment of tumour response to treatment: comparison of linear, cross-sectional and volumetric measures of tumour size. *Br J Radiol* 2000; **73**: 1178–84.
4. Werner-Wasik M, Xiao Y, Pequignot E, Curran WJ, Hauck W. Assessment of lung cancer response after nonoperative therapy: tumor diameter, bidimensional product, and volume. A serial CT scan-based study. *Int J Radiat Oncol Biol Phys* 2001; **51**: 56–61.
5. Warren KE, Patronas N, Aikin AA, Albert PS, Balis FM. Comparison of one-, two-, and three-dimensional measurements of childhood brain tumors. *J Natl Cancer Inst* 2001; **93**: 1401–5.
6. James K, Eisenhauer E, Christian M, Terenziani M, Vena D, Muldal A, Therasse P. Measuring response in solid tumors: unidimensional versus bidimensional measurement. *J Natl Cancer Inst* 1999; **91**: 523–8.
7. Sobin LH, Wittekind Ch., editors. UICC TNM Classification of malignant tumours. 5th ed. New York: John Wiley & Sons Inc; 1997.
8. Breiman RS, Beck JW, Korobkin M, Glenny R, Akwari OE, Heaston DK, Moore AV, Ram PC. Volume determinations using computed tomography. *Am J Roentgenol* 1982; **138**: 329–33.
9. Van Hoe L, Van Cutsem E, Vergote I, Baert AL, Bellon E, Dupont P, Marchal G. Size quantification of liver metastases in patients undergoing cancer treatment: reproducibility of one-, two-, and three-dimensional measurements determined with spiral CT. *Radiology* 1997; **202**: 671–5.
10. Bozcuk HS, Ravi R, Turner B, Tsetis D, Thomas JM, Chan O, Reznek R, Hendry WF, Oliver RT. Computed tomography 21 days after chemotherapy, three-dimensional estimates of metastatic volume and the need for surgery in patients with germ cell cancer. *BJU Int* 2000; **86**: 707–13.