CI

e-med
www.e-med.org.uk

ARTICLE

Response to chemotherapy in patients with lung metastases: how many nodules should be measured?

R Chojniak*, L S Yu* and R N Younes†

Departments of *Radiology; †Thoracic Surgery, Hospital A C Camargo, LIM-62, UNIP, São Paulo, Brazil

Corresponding address: Rubens Chojniak, MD PhD, Department of Radiology, Hospital do Cancer A C Camargo, rua Prof. Antonio Prudente 211, Liberdade, São Paulo, Brazil 01509-010. E-mail: chojniak@uol.com.br

Abstract

Objective: With the introduction of cross-sectional imaging methods the number of lesions per patient that can be evaluated is frequently large and most oncologists and study protocols use only one lesion or a few 'representative' lesions to evaluate chemotherapy response. Intra-patient response variability can therefore affect evaluation reproducibility. This study evaluates intra-individual variation in response to chemotherapy in patients with multiple lung metastases.

Methods: We prospectively studied chest CT images of patients with solid tumors and pulmonary metastases under systemic chemotherapy being evaluated for tumor response. The response of 566 pulmonary nodules in 41 evaluations was determined by both WHO and RECIST criteria in order to determine intra-individual tumor response variation.

Results: There was almost perfect agreement between the WHO and the RECIST criteria for the evaluation of tumor response. High intra-individual variability of tumor response was observed in a significant proportion of the evaluations. A new nodule was the main criterion for determination of disease progression. A mean of 35% of the total number of nodules of a patient have a response evaluation different from that calculated with all the nodules together.

Conclusions: Intra-individual variation in tumor response of pulmonary metastases is elevated in some patients. Selecting any or some nodules for response evaluation could significantly influence therapeutic response perception.

Keywords: *Antineoplastic chemotherapy; protocols Phase II; FDA drug evaluation; computed tomography.*

Introduction

Systemic chemotherapy is the main treatment modality in the management of patients with metastatic disease[1]. After chemotherapy, oncologists evaluate tumor response by observing tumor behavior, i.e. growth, reduction or stability in its dimensions[2,3]. Nowadays, tumor response to therapy as determined by imaging methods is generally used to inform decisions regarding either maintenance or interruption of treatment. In the late 1970s the World Health Organization (WHO) introduced a standard assessment of tumor response, proposed by Miller *et al.*[2] and adopted internationally, defining objective responses of lesions measurable in two dimensions,

such as pulmonary metastasis assessed by X-rays. This evaluation is performed by comparison of tumor area, the sum of all lesions' greater perpendicular diameter products, measured in a planar image. More recently a new set of guidelines has been introduced by the Response Evaluation Criteria in Solid Tumors (RECIST) group with the purpose of reviewing the former criteria and better standardizing response evaluation[3]. This model uses a unidimensional approach taking the sum of the longest diameters instead of the sum of the areas.

With the introduction of cross-sectional imaging methods, the number of measurable metastatic lesions detected in a single patient has increased dramatically, and most oncologists (as well as study protocols)

recommend the use of one lesion or a few 'representative' lesions to evaluate response in individual patients with multiple lesions[4]. Metastatic nodules are not uniform and consist of a heterogeneous cell population with diverse biological behavior that could account for differences in chemotherapy response[5]. A wide range of growth patterns in pulmonary metastases of patients not previously submitted to treatment has been observed[6]. This variation in behavior, if observed in patients being evaluated for chemotherapy response, could influence the response perception. The selection of one or a few nodules, instead of including all identifiable nodules in the response evaluation, could lead to misevaluation and consequently to the continuity of ineffective treatment or the interruption of potentially effective therapy.

The present study used CT to quantify the variation in tumor response of pulmonary metastases of solid tumors of varied histology in individual consecutive patients that were submitted to CT studies in order to evaluate response to chemotherapy. We submitted each pulmonary nodule individually, as if it were a solitary metastasis, as well as all the nodules of the same patient combined, to both the WHO and the RECIST criteria and compared the response evaluations in each setting.

Patients and methods

We prospectively evaluated two consecutive chest CT scans of patients with the diagnosis of solid tumor and pulmonary metastases receiving systemic chemotherapy and being routinely evaluated for tumor response. Informed consent was obtained from all the patients. We included in this study 41 chemotherapy response evaluations in 33 patients (20 women and 13 men), with ages ranging from 14 to 81 years (median 46 years). In eight patients a second response evaluation was obtained. The types of primary cancer are listed in Table 1. The attending physician (Department of Clinical Oncology) established the diagnosis of pulmonary metastases usually by the presence of new pulmonary nodules and progression of metastatic disease and the type of chemotherapy used. The interval between the CT evaluations varied from 1.25 to 8.8 months (median: 3.9;

Table 1 Distribution of primary types of metastatic solid tumor

Primary tumor	n
Breast cancer	13
Bone or soft tissue sarcoma	9
Testicular cancer	3
Kidney cancer	2
Colon cancer	2
Stomach cancer	1
Lung cancer	1
Cervix cancer	1
Bladder cancer	1
Total	33

mean ($\pm$sd) 3.8 ($\pm$1.6) months) and the number of cycles varied from 2.0 to 6.0 cycles (median: 4; mean $\pm$ sd: 3.7 $\pm$ 1.2 cycles). No patient received radiotherapy between the two CT scan examinations.

Helical scan techniques were performed on CT Prospeed and CT Hispeed scanners (General Electric). The slices obtained were contiguous with 7 mm thickness and a pitch of 1.5 or less. All nodules identified in each CT examination with clear margins were numbered and measured. The same radiologist took measurements with a caliper on hard copies. The two larger perpendicular diameters on the axial plane were measured in images printed in lung windows. We studied 566 pulmonary nodules. The number of nodules in an individual patient varied from 2 to 69 (median: 7; mean $\pm$ sd: 13.8 $\pm$ 15.0 nodules). The nodules' initial larger diameter ranged from 2 to 82 mm (median: 10; mean $\pm$ sd: 11.6 $\pm$ 8.5 mm).

In each set of CT scan examinations performed to assess tumor response to chemotherapy, WHO and RECIST criteria classified each nodule individually. For the patient's response evaluation a modified version of the WHO and RECIST criteria was used by considering the sum of measurements of all pulmonary nodules in each patient to better represent the total tumor volume change for each patient.

The bidimensional WHO criteria of tumor response categories are: (a) partial response when area reduction is 50% or more; (b) stable disease for a reduction of less than 50% or an increase less than 25%; and (c) disease progression for an increase of 25% or more. The unidimensional RECIST criteria of tumor response categories are: (a) partial response if linear larger dimension reduction is 30% or more; (b) stable disease for a reduction less than 30% or an increase less than 20%; and (c) disease progression for an increase of 20% or more. The disappearance of the lesion(s) is considered a complete response by both assessment criteria and both consider the presence of any new nodule as progression of disease, independent of the behavior of any other nodule.

The following parameters were determined: (a) individual nodule response rate evaluation; (b) response evaluation for each patient, as a whole, according to the WHO and the RECIST criteria; (c) intra-individual distribution of response for every patient's nodule; and (d) the proportion of nodules evaluated differently from the patient's response, taking the sum of all nodules into consideration. The nodules' response category distributions for different types of cancer, chemotherapy, and initial number of nodules were compared by Chi-square test. Differences were considered statistically significant for $P < 0.05$. The agreement between both response evaluation criteria was calculated by the Kappa-interrater agreement index[7].
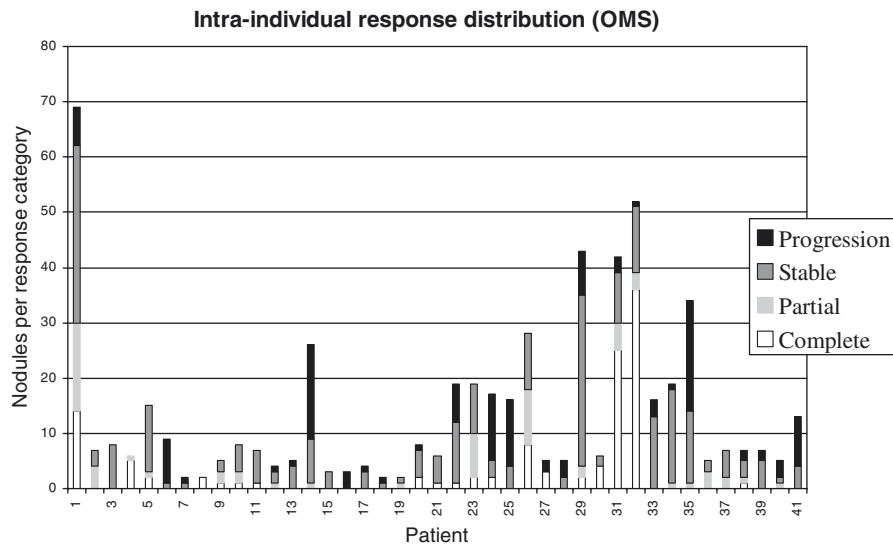
Figure 1 **Intra-individual distribution of response evaluation of pulmonary metastases by World Health Organization criteria.**

Results

All nodules ($n = 566$) were evaluated. Half ($n = 283$) of the nodules showed a reduction in size considering only the larger measured diameter, while 126 (22%) remained unaltered comparing both CT scan studies and 157 (28%) increased in size. According to the WHO criteria, 113 (20%) had a complete response, 66 (12%) had a partial response, 258 (46%) were stable and 129 (23%) progressed. Using the RECIST criteria, 134 nodules were considered measurable as having a diameter twice the size of the slice thickness utilized (>14 mm). Of these, 22 (16%) nodules had a complete response, 17 (13%) a partial response, 67 (50%) were stable and 28 (21%) progressed. Classification was different by the two criteria in 13 (10%) of the 134 measurable nodules. The Kappa interrater agreement for both criteria evaluations was 0.85 (Table 2).

Table 2 *Response evaluation of measurable nodules ($>$14 mm) assessed by the WHO and the RECIST criteria*

WHO	RECIST				
	Complete	Partial	Stable	Progression	Total
Complete	22	0	0	0	22
Partial	0	15	2	0	17
Stable	0	2	56	1	59
Progression	0	0	9	27	36
Total	22	17	67	28	134

Kappa index: 0.85.

Evaluating the patients' response, taking into account the sum of all nodules, by the WHO criteria, one patient was classified as having complete response, five partial response, 15 stable disease and 20 progressive disease; and by the RECIST criteria one patient was classified as having complete response, five partial response, 16 stable disease and 19 progressive disease. Three patients had different response evaluations by the different criteria. The Kappa interrater agreement for both criteria evaluations was 0.90 (Table 3).

Table 3 *Response evaluation of patients assessed by the WHO and the RECIST criteria*

WHO	RECIST				
	Complete	Partial	Stable	Progression	Total
Complete	1	0	0	0	1
Partial	0	4	1	0	5
Stable	0	1	14	0	15
Progression	0	0	1	19	20
Total	1	5	16	19	41

Kappa index: 0.90.

The intra-individual variation for metastases response evaluation was quite diverse by both criteria, with no relation to type of cancer, chemotherapy or number of nodules. By the WHO criteria: from the total of 41 response evaluations, all the patients' nodules had the same response classification in four; in 16 evaluations there were nodules in two distinct classifications; in 12 three different classifications; and in seven, nodules in the four possible categories (Fig. 1).

By the RECIST criteria: measurable nodules were present in 33 of the 41 response evaluations; in seven there was only one nodule; in 11 all the patients' nodules had the same response classification; in 12 evaluations there were nodules in two distinct classifications; and in three evaluations there were nodules in three different classifications (Fig. 2).

Eighteen patients presented 82 new nodules observed at the second CT scan, varying from 1 to 21 per patient (median 2, mean 4.8), in a proportion of the initial total
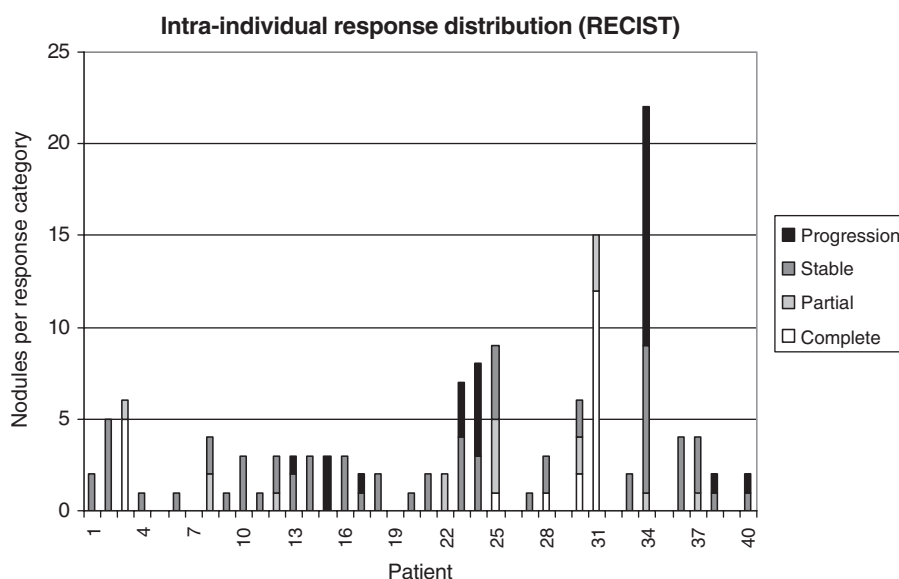
Intra-individual response distribution (RECIST)

Figure 2   **Intra-individual distribution of response evaluation of pulmonary metastases by the Response Evaluation Criteria in Solid Tumors (RECIST) group criteria.**

number of nodules ranging from 5% to 525% (median 25%, mean 75%). By the RECIST criteria, in one of the 19 patient evaluations resulting in disease progression no new nodule was present and in 14 a new nodule was the sole criterion for disease progression classification (Table 4).

By the RECIST criteria, the proportion of measurable nodules classified differently from the patient evaluation, taking the sum of all measurable nodules into consideration, varied from 0% to 100% (median: 25%; mean ± sd: 30% ± 33%). In a patient with three lung metastases, two of them were evaluated as stable and one as complete response. The patient evaluation by comparison of the sum of diameters of the three nodules resulted in partial response. In 15 evaluations there were measurable lesions and there were no new nodules; when the measurable response was the only considered factor, the proportion of measurable nodules classified differently from the patient evaluation, taking the sum of all measurable nodules into consideration, varied from 0% to 100% (median: 33%; mean ± sd: 35% ± 35%).

Discussion

After chemotherapy, evaluation of tumor response is obtained by observing the progress of lesion size[1]. Difficulties arose when the number of lesions per patient that could be evaluated increased, mainly after the introduction of cross-sectional imaging methods[4]. The assessment of tumor response for every single lesion became time-consuming. Nowadays most radiologists and oncologists use only one lesion or a few 'representative' lesions to evaluate tumor response in patients presenting with multiple nodules[3]. Thus, it is important

to know the intra-individual variability of response rate evaluation for different tumors in the same patient.

Clinically, the response rate of solid tumors has been calculated by taking the tumor diameters in observations separated by the treatment, and by determining tumor shrinkage, stability or growth. For many years the World Health Organization criteria for treatment response evaluation have been the criteria used by most oncologists and in most clinical trials[2]. More recently, several organizations involved in clinical research have reviewed these criteria on the basis of experience acquired since they were introduced and a new set of guidelines has been developed with a model by which response rates could be derived from unidimensional measurement lesions instead of the former bidimensional approach[3]. According to this model, the use of only one lesion dimension simplifies the task of evaluating tumor response and correlates well with the lesion's area, previously used. The RECIST group tested their criteria in several historical study protocols and obtained good correlation with the WHO criteria. More recently, studies have shown both good and poor correlation between the unidimensional and bidimensional response evaluations[8,9]. In this present study a high correlation between the WHO and the RECIST criteria in the evaluation of individual nodules and patients was achieved. The advantage of using the RECIST criteria is the simplicity of taking only one measurement per lesion.

Nowadays, the best method of assessing tumor response in most clinical situations is CT scan, leading to detection of a larger number of lesions, with smaller diameters, and more precision on measuring when compared to other methods[10,11]. Even though pulmonary metastases evaluated by CT scan should be

one of the best scenarios in terms of measurement of lesions, since the low-density lung provides natural contrast for the dense pulmonary nodule, variability in the application of response criteria could compromise the reproducibility of results[12]. Major potential sources of response evaluation variation are imaging techniques, inter-observer variability, and the selection of target lesions[12,13]. In order to minimize observer variation, this study applied helical CT scanning and had the lesions measured by the same radiologist to evaluate response of the multiple metastases in the same patient, as routinely done in the clinical setting and in study protocols[13].

Table 4   Impact of the presence of new nodules in the patient's global response evaluation assessed by the RECIST criteria

	Response of sum of measurable nodules	New nodule	Patient response
1	Stable	–	Stable
2	**Stable**	+	**Progression**
3	Stable	–	Stable
4	Partial	–	Partial
5	Stable	–	Stable
6	**Stable**	+	**Progression**
7	Stable	–	Stable
8	Complete	–	Complete
9	Partial	–	Partial
10	Stable	–	Stable
11	Stable	–	Stable
12	**Stable**	+	**Progression**
13	Stable	–	Stable
14	**Stable**	+	**Progression**
15	Stable	–	Stable
16	Progression	+	Progression
17	**Stable**	+	**Progression**
18	Stable	–	Stable
19	Stable	–	Stable
20	**Stable**	+	**Progression**
21	Stable	–	Stable
22	**Stable**	+	**Progression**
23	**Partial**	+	**Progression**
24	Progression	+	Progression
25	Progression	+	Progression
26	Partial	–	Partial
27	**Stable**	+	**Progression**
28	**Stable**	+	**Progression**
29	Complete	–	Partial
30	**Stable**	+	**Progression**
31	**Partial**	+	**Progression**
32	Partial	–	Partial
33	Stable	–	Stable
34	Stable	–	Stable
35	Progression	–	Progression
36	Stable	–	Stable
37	Stable	–	Stable
38	Stable	–	Stable
39	**Stable**	+	**Progression**
40	**Stable**	+	**Progression**
41	Progression	+	Progression

Bold type indicates evaluations in which a new nodule was the sole criterion for disease progression.

The presence of a new nodule was the main factor for determining disease progression, and the proportion of these new nodules in some cases was as low as 5%. The systematic evaluation of all the patients' nodules is then necessary to guarantee the identification of a new nodule.

We observed a large variation in tumor response for lesions in the same patient—i.e. the individual nodules in the same patient have individual and unpredictable behavior and can be evaluated as independent lesions. By using the WHO and the RECIST criteria of tumor response to therapy we verified that there were nodules labeled as disease progression in patients classified as stable, and unaltered nodules in patients evaluated as having progressive disease. In our patient population the proportion of nodules in a patient with response evaluation different from that obtained taking the sum of all nodules together amounted to 35%.

The results confirmed our previous impression that the application of the WHO or the RECIST method to one or some nodules in a patient, as frequently utilized by protocols, could also be misleading. In the same patient, one could select a stable nodule, a nodule with growth or a nodule in regression, and this could result in a decision error regarding the treatment regimen. The use of the sum of the evaluation of all nodules together maximizes the reproducibility of response evaluations.

In this study we tried to reproduce what seem to be the usual conditions for tumor response evaluation in most protocols and institutions; some of these conditions are limitations for the best possible response evaluation accuracy, but are close to clinical practice standard conditions. There was no histological proof of the metastatic nature of the lung nodules, and some could have been infectious or inflammatory; in clinical practice, however, histological proof is rarely required and the presence of new pulmonary nodules and overall oncologic disease progression are the criteria adopted by the oncologist for a diagnosis of lung metastases. When this study was conducted our CT protocol for lung metastases evaluation was a slice thickness of 7 mm. Nowadays many institutions with faster and multiple detector CTs adopt 5 mm or less in the evaluation of lung metastases. However, the recommendation of the authors of the RECIST group criteria was followed in measuring lesions no smaller than twice the size of the slice thickness[3]. Hard copy measurements were taken, despite having been shown to be less reproducible than computer display methods, as computer measurement is not always available, especially to the oncologist.

In conclusion, there is a high correlation between the World Health Organization criteria and the RECIST group criteria. Intra-individual variation in tumor response of pulmonary metastases is elevated in some patients, and chemotherapy response evaluation utilizing only one or some of the patient's nodules could lead to inappropriate reproducibility in therapeutic response perception.

References

[1] Chu E, DeVita VT Jr. Chemotherapy as part of the initial treatment of cancer. In: Cancer: Principles and Practice of Oncology, 6th edn. DeVita VT Jr, Hellman S, Rosenberg SA, eds. Philadelphia, PA: Lippincott, Williams & Wilkins, 2001: 290.

[2] Miller AB, Hoogstraten B, Staquet M, Winkler A. Reporting results of cancer treatment. Cancer 1981; 47: 207–14.

[3] Therasse P, Arbuck SG, Eisenhauer EA, Wanders J, Kaplan RS, Rubinstein L *et al*. New guidelines to evaluate the response to treatment in solid tumors. J Natl Cancer Inst 2000; 92: 205–16.

[4] Davis SD. CT evaluation for pulmonary metastases in patients with extrathoracic malignancy. Radiology 1991; 180: 1–12.

[5] Fidler IJ, Poste G. The cellular heterogeneity of malignant neoplasms: implications for adjuvant chemotherapy. Semin Oncol 1985; 12: 207–21.

[6] Chojniak R, Younes RN. Pulmonary metastases tumor doubling time: assessment by computed tomography. Am J Clin Oncol 2003; 26: 374–7.

[7] Altman DG. Practical Statistics for Medical Research, Inter-rater agreement. London: Chapman & Hall, 1991: 403–9.

[8] Prasad SR, Saini S, Sumner JE, Hahn PF, Sahani D, Boland GW. Radiological measurement of breast cancer metastases to lung and liver: comparison between WHO (bidimensional) and RECIST (unidimensional) guidelines. J Comput Assist Tomogr 2003; 27: 380–4.

[9] Tran LN, Brown MS, Goldin JG, Yan X, Pais RC, McNitt-Gray MF *et al*. Comparison of treatment response classifications between unidimensional, bidimensional, and volumetric measurements of metastatic lung lesions on chest computed tomography. Acad Radiol 2004; 11: 1355–60.

[10] Gurland J, Johnson RO. How reliable are tumor measurements? JAMA 1965; 194: 125–30.

[11] Erasmus JJ, Gladish GW, Broemeling L, Sabloff BS, Truong MT, Herbst RS *et al*. Interobserver and intraobserver variability in measurement of non-small-cell carcinoma lung lesions: implications for assessment of tumor response. J Clin Oncol 2003; 21: 2574–82.

[12] Thiesse P, Ollivier L, Di Stefano-Louineau D, Negrier S, Savary J, Pignard K *et al*. Response rate accuracy in oncologic trials: reasons for interobserver variability. J Clin Oncol 1997; 15: 3507–14.

[13] Bogot NR, Kazerooni EA, Kelly AM, Quint LE, Desjardins B, Nan B. Interobserver and intraobserver variability in the assessment of pulmonary nodule size on CT using film and computer display methods. Acad Radiol 2005; 12: 948–56.