

Comparison of WHO and RECIST Criteria for Response in Metastatic Colorectal Carcinoma

Jung-Hye Choi, M.D.¹, Myung-Ju Ahn, M.D.¹, Hyan-Chul Rhim, M.D.², Jin-Woo Kim, M.D.², Gang-Hong Lee, M.D.³, Young-Yeul Lee, M.D.¹ and In-Soon Kim, M.D.¹

Departments of ¹Internal Medicine, ²Diagnostic Radiology and ³Surgery, Hanyang University College of Medicine, Seoul, Korea

Purpose: This study compared the WHO criteria with the response evaluation criteria in solid tumors (RECIST) in the same patients with metastatic colorectal cancer in order to determine the significance of the RECIST. In addition, this study compared the estimations of medical oncologists with those of a radiologist.

Materials and Methods: Between 2002 and 2005, a total of 48 patients (male: female ratio, 29:19; median age, 58 years) with measurable lesions receiving chemotherapy for metastatic colorectal carcinoma were enrolled in this study. Two medical oncologists and one radiologist, who were blinded to the patients' condition, independently reviewed all the CT images. The results were compared using a kappa test.

Results: The kappa test for concordance between the WHO and RECIST criteria of the medical oncologists and the radiologist were 0.908 and 0.841, respectively. The level of concordance between the investigators using the WHO and RECIST were 0.722 and 0.753, respectively.

Conclusion: The RECIST criteria are comparable to the WHO criteria in evaluating the response of colorectal carcinoma, but have simple and reproducible guidelines. The use of RECIST is recommended for evaluating the treatment efficacy in clinical trials and practice. (*Cancer Res Treat. 2005;37:290-293*)

Key Words: Tumor response, RECIST, WHO, Metastatic colorectal carcinoma

INTRODUCTION

An evaluation of the tumor response to treatment is important in clinical trials for new drugs as well as the routine management of advanced malignancies. Since the early 1980s, the WHO response criteria were adopted as the standard method for evaluating the tumor response (1). According to the WHO criteria, the total tumor size is determined by bidimensional measurements e.g. the sum of the products of the two longest diameters in the perpendicular dimensions of all tumors. The tumor response to treatment is divided into four categories. However, some problems have arisen when using these criteria and a new methodology is required (2). Recently, the RECIST (Response Evaluation Criteria in Solid Tumors) was proposed as a new guideline for evaluating the response using unidimensional measurements instead of bidimensional measurements, a lower number of measured lesions, the withdrawal of the progression criteria based on the isolated increase in a single lesion, and a different shrinkage threshold for defining the tumor response and progression (3). In order to clarify the

significance of these new guidelines, the WHO and RECIST criteria were compared in the same patients with metastatic colorectal cancer, and the estimations by medical oncologists were compared with those of a radiologist.

MATERIALS AND METHODS

1) Patients

Between March 2002 and March 2005, a total of 48 patients, who received chemotherapy as the first or second line treatment for metastatic colorectal carcinoma at Hanyang University Hospital, were enrolled in this study. There were 29 males (60.4%) and 19 females (39.6%), with a median age of 58.0 years (range: 31~76 years). Thirty-five patients were treated with FOLFIRI (irinotecan, leucovorin, 5-fluorouracil), nine patients were treated with XELOX (capecitabine, oxaliplatin), and four patients received FOLFOX (oxaliplatin, leucovorin, 5-fluorouracil). All the patients had at least one bidimensionally measurable lesion > 10×10 mm, as assessed by the CT scan.

2) Tumor measurement

The CT scans were obtained at the metastatic sites using a standardized contrast enhanced imaging protocol. All the CT scans used a spiral CT (Siemens Medical Systems, Erlangen, Germany). The CT image data were reconstructed with a 5 mm thickness and were displayed on the monitors of a picture archiving and communications system (PACS). The tumor size was measured from digitalized images with electronic calipers. The tumor response was assessed using the WHO and RECIST

Correspondence: Myung-Ju Ahn, Department of Internal Medicine, Hanyang University College of Medicine, 17 Haengdang-dong, Seongdong-gu, Seoul 133-792, Korea. (Tel) 82-2-2290-8335, (Fax) 82-2-2298-9183, (E-mail) ahnmj@hanyang.ac.kr

Received July 11, 2005, Accepted August 26, 2005

This work was supported by the research fund (2002) of Hanyang University.

Table 1. The WHO and RECIST* criteria

	WHO	RECIST*
Method	Sum of products of the two longest diameters in perpendicular dimensions (bidimensional)	Sum of longest diameters (unidimensional)
No. of measured lesion	All lesions	Target lesion: maximum 5 per organ, 10 in total
Response		
CR [†]	Disappearance of all known disease, confirmed at 4 wks	Disappearance of all known disease, confirmed at 4 wks
PR [‡]	≥ 50% decrease, confirmed at 4 wks	≥ 30% decrease, confirmed at 4 wks
SD [§]	Neither PR nor PD criteria met	Neither PR nor PD criteria met
PD	≥ 25% increase, no CR, PR, or SD documented before increased disease, new lesion (s), ≥ 25% increase in 1 lesion	≥ 20% increase, no CR, PR, or SD documented before increased disease, new lesion (s)

*response evaluation criteria in solid tumors, [†] complete response, [‡] partial response, [§] stable disease, ^{||} progressive disease.

criteria. Table 1 gives the definition of the best response according to the WHO or RECIST criteria. Two medical oncologists reviewed CT images together and one radiologist separately reviewed all the CT images without any information of the patients. The tumor measurements were performed two times for each lesion. The results were compared using the kappa test for the concordance of the response. If the kappa value was > 0.8, the concordance was considered to be excellent and a kappa value ranging from 0.60 to 0.80 was considered to be good.

RESULTS

Table 2 and Fig. 1 show a comparison of the WHO and RECIST criteria estimated by the two medical oncologists and the radiologist. When the medical oncologists performed the tumor measurements, the overall response and progression rates according to the WHO criteria were 35.4% (1 CR and 16 PR) and 29.2% (14 PD), respectively. Using the RECIST criteria, two patients with SD based on the WHO criteria were reclassified as PR, and one patient with PR was downgraded to SD. The overall response rate was then 37.5% (1 CR, 17 PR). The kappa test for the concordance of the medical oncologists using the WHO and RECIST criteria was 0.908. When the radiologist performed the tumor measurements, there were 1 CR, 14 PR, 21 SD, and 12 PD according to the WHO criteria, and 1 CR, 15 PR, 22 SD, and 10 PD according to the RECIST criteria. The overall response rate according to the WHO criteria was 31.3%. Five patients were reclassified using the RECIST criteria, and the overall response rate was 33.3%. The kappa test for the concordance of the overall response was 0.841.

The treatment response using WHO were concordant in 39 patients (81.3%) between the two investigator groups, and those using the RECIST criteria were concordant in 40 patients (83.3%). The level of concordance between the investigators using the WHO and RECIST were 0.722 and 0.753, respectively.

Table 2. Comparison of the WHO and RECIST criteria estimated by the medical oncologist (A) and radiologist (B).

A					
RECIST criteria	WHO criteria				
	CR	PR	SD	PD	Total
CR	1	0	0	0	1
PR	0	15	2	0	17
SD	0	1	15	0	16
PD	0	0	0	14	14
Total	1	16	17	14	48

B					
RECIST criteria	WHO criteria				
	CR	PR	SD	PD	Total
CR	1	0	0	0	1
PR	0	13	2	0	15
SD	0	1	19	2	22
PD	0	0	0	10	10
Total	1	14	21	12	48

DISCUSSION

A standardized approach to measuring a tumor and determining the response criteria is important for making appropriate

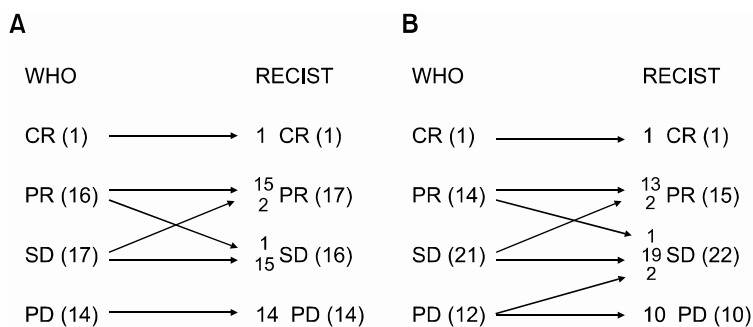


Fig. 1. Change in the response between the WHO and RECIST criteria estimated by the medical oncologist (A) and radiologist (B).

medical decisions. In order to prevent an inappropriate designation of a tumor response, the tumor measurements should be standardized and be accurate, with a low intra-and inter-observer variability. For more than 2 decades, the WHO response criteria have been standard method for making a radiological tumor evaluation. However, these criteria are quite complex. Measuring all visible lesions in two dimensions is considerably time-consuming and has a risk of error. In 1994, a large international working group was established to review the guidelines. After several years of intensive discussion and an analysis of 14 large clinical trials, which included more than 4,500 patients, the working group concluded that unidimensional tumor measurements provide results equivalent to those obtained using the bidimensional criteria and finally recommended the simpler new guidelines, the RECIST criteria (3).

The theoretical background for the RECIST criteria is that the sum of the largest diameters of the individual tumors is more linearly related to the level of cell death than the sum of the bidimensional products (4). In the RECIST criteria, a PR is defined as at least a 30% reduction in the sum of the longest diameter of the target lesions, and a PD is defined as at least a 20% increase in the sum of the longest diameter. Assuming that tumors are spherical, a 30% decrease in the sum of the diameters, which corresponds to a 65% decrease in the tumor volume, is equivalent to a 50% reduction in the sum of the bidimensional products. Accordingly, the threshold of a PR using the two criteria is almost identical. However, the limit of PD is higher in the RECIST criteria, a 20% increase in the largest diameter represents a 73% increase in the tumor volume versus a 40% increase in the tumor volume when there is a 25% increase in the bidimensional product. One of the new concepts of the RECIST is the documentation of the target and non-target lesions. A maximum of five lesions per organ and 10 lesions in total should be identified as target lesions, and the sum of the longest diameter for all the target lesions will be calculated.

After the proposal of the RECIST criteria, several studies were carried out to validate this new guideline in various solid tumors (5~7). These analyses concluded that the RECIST criteria are as effective as the WHO criteria in terms of the response rate. However, inter-observer and intra-observer variations are still an unsolved problem in the RECIST criteria, and this method is not suitable for several tumors such as a pleural mesothelioma (8). Erasmus et al reported that the measurements of the lung tumor size on a CT scan using bidimensional and unidimensional methods are often inconsistent. This can lead

to an incorrect interpretation of the tumor response and a greater inter-observer variability than intra-observer variability. Moreover, the measurement differences are greatest when the edge of the lesion is irregular or speculated (9).

This study showed excellent agreement between the WHO and RECIST criteria estimated by two medical oncologists and one radiologist, respectively. The results of the kappa test for concordance in the overall response were > 0.8 . A discordant result is usually obtained when a tumor has an irregular or asymmetric shape or when the calculated sum is close to the threshold level. There might be discrepancies in measuring the tumor response rate between an oncologist and radiologist due to errors in the tumor measurements and errors in selecting the target lesions. It is believed that the inter-observer variability using the RECIST criteria may be lower than when using the WHO criteria because the RECIST criteria are simpler to apply. Therefore, the estimations of medical oncologists were compared with those of a radiologist. However, both results using the RECIST and WHO criteria showed good agreement between the two investigator groups. Therefore, it is essential that medical oncologists and radiologists work in unison when evaluating the tumor response.

CONCLUSIONS

These results suggest that the RECIST criteria are comparable to the WHO criteria in evaluating the response in a colorectal carcinoma using simple and reproducible guidelines. Although the use of the RECIST criteria is useful for evaluating the treatment efficacy in clinical trials and practice, the limitation of the RECIST criteria highlight the need for additional response analysis techniques.

REFERENCES

1. Miller AB, Hoogstraten B, Staquet M, Winkler A. Reporting results of cancer treatment. *Cancer*. 1981;47:207-14.
2. Thiesse P, Ollivier L, Di Stefano-Louineau D, Negrier S, Savary J, Pignard K, et al. Response rate accuracy in oncology trials: reasons for interobserver variability. *Groupe Francais d'Immunotherapie de la Federation Nationale des Centres de Lutte Contre le Cancer. J Clin Oncol*. 1997;15:3507-14.
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. *European Organization for*

- Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J Natl Cancer Inst.* 2000;92:205-16.
4. James K, Eisenhauer E, Christian M, Terenziani M, Vena D, Muldal A, et al. Unidimensional versus bidimensional measurement. *J Natl Cancer Inst.* 1999;91:523-8.
 5. Trillet-Lenoir V, Freyer G, Kaemmerlen P, Fond A, Pellet O, Lombard-Bohas C, et al. Assessment of tumour response to chemotherapy for metastatic colorectal cancer: accuracy of the RECIST criteria. *Br J Radiol.* 2002;75:903-8.
 6. Park JO, Lee SI, Song SY, Kim K, Kim WS, Jung CW, et al. Measuring response in solid tumors: comparison of RECIST and WHO response criteria. *Jpn J Clin Oncol.* 2003;33:533-7.
 7. 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.
 8. Byrne MJ, Nowak AK. Modified RECIST criteria for assessment of response in malignant pleural mesothelioma. *Ann Oncol.* 2004;15:257-60.
 9. 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.
-