New Evaluation Criteria for Response and Toxicity in Lung Cancer Treatment

JMAJ 46(12): 554–558, 2003

Masahiko SHIBUYA
Chief, Division of Respiratory Medicine, Tokyo Metropolitan Komagome Hospital

Abstract: Response Evaluation Criteria in Solid Tumors (RECIST) and the National Cancer Institute—Common Toxicity Criteria (NCI-CTC) Version 2.0 were prepared for the purpose of standardizing the evaluation criteria for response and toxicity of cancer treatment. Objective response evaluation in RECIST has been simplified from the method using bidimensional measurement (product of 2 diameters) in WHO Standards to the summation of unidimensional measurement (largest diameter). NCI-CTC Version 2.0 includes toxicities and categories that have been increased from 49 toxicities in 18 categories in the previous version to 279 toxicities in 24 categories. It also provides more appropriate toxicity evaluation by grading of associated toxicities.

Key words: Response evaluation criteria; Toxicity evaluation criteria; RECIST; NCI-CTC Version 2.0

Introduction

The clinical effectiveness of various therapies for malignant tumors is measured in terms of objective response (tumor size reduction) assessed shortly after treatment and the assessment of long-term results, which include the time from treatment to relapse or progression and the survival time.

In 1979, the World Health Organization (WHO) published the “WHO Handbook for Reporting Results of Cancer Treatment” stipulating the standards for summary reporting of objective response and long-term results (hereinafter called “WHO Standards”). The WHO Standards have been used internationally, and Japanese physicians are using the Japan Society of Clinical Oncology Clinical Response Evaluation Criteria for Chemotherapy in Solid Tumor (RECIST) formulated based on the WHO Standards.

However, various problems in the WHO Standards have been pointed out, and various groups have modified the standards. To ensure international harmonization, a revised edition of the WHO Standards called the Response Evaluation Criteria in Solid Tumors (RECIST) was prepared in 1999.

Because the usefulness of cancer chemotherapy is evaluated based on the comprehen-
sive assessment of anti-tumor effects and toxicity, the evaluation of toxicity is as important as the evaluation of anti-tumor effects.

The method of toxicity evaluation in clinical trials has conventionally been based on the “WHO Criteria for the Evaluation of Side Effects”,1) National Cancer Institute — Common Toxicity Criteria (NCI-CTC), and other references. A revised version of NCI-CTC was prepared in 1998 in response to the need for internationally harmonized toxicity evaluation criteria based on the new GCP (Good Clinical Practice: Ministry of Health and Welfare Ordinance on Standards for the Implementation of Clinical Trials on Pharmaceutical Products). At present, this new NCI-CTC is used internationally as the criteria for toxicity evaluation.

This article explains the outline of RECIST and the features of the revised NCI-CTC (Version 2.0).

Outline of New Response Evaluation Criteria (RECIST)

1. Measurement of tumors

While current WHO Standards use bidimensional measurement (the product of the length of the major axis and the longest diameter intersecting with it at right angles), the new RECIST uses unidimensional measurement because of the theoretical consideration that unidimensional measurement is in better proportion to the number of tumor cells than bidimensional measurement and because of the ease of measurement.

2. Definition of lesions measured as tumors at baseline

(1) Measurable lesions

These are defined as lesions that can be measured unidimensionally and have a longest diameter of 20 mm or more on conventional CT scan and 10 mm or more on helical CT scan. This definition includes only lesions with a length of the major axis at least twice as large as the slice thickness.

(2) Non-measurable lesions

These lesions include small lesions with a longest diameter of <20 mm on conventional CT and <10 mm on helical CT, as well as true non-measurable lesions [bone lesions, carcinomatous meningitis, ascites, pleural effusion, pericardial effusion, inflammatory breast lesions, lymphangitis (skin/lung), abdominal tumors that cannot be identified and traced by imaging techniques, and cystic lesions].

3. Methods of measurement

(1) Clinical measurement (external measurement)

Only superficial lesions, such as skin tumors and palpable lymph nodes, are considered measurable. In the case of skin lesions, it is recommended to use color photographs including a ruler to indicate lesion size.

(2) Chest X-ray

Lesions on chest X-ray are acceptable as measurable lesions when they are clearly defined and surrounded by aerated lung. However, CT is preferable.

(3) CT and MRI

CT and MRI are the best currently available and reproducible methods for measuring target lesions selected for response assessment. Conventional CT and MRI in thoracic, abdominal, and pelvic regions should be performed with cuts of 10 mm or less in slice thickness. Helical CT in these regions should be performed using a 5 mm reconstruction algorithm. Head and neck tumors and those of extremities usually require specific protocols.

(4) Ultrasound

When the primary endpoint of the study is objective response evaluation, ultrasound should not be used to measure tumor lesions that are not directly measurable from the body surface. Ultrasound is, however, a possible alternative to clinical measurements of superficial palpable superficial nodules, subcutaneous lesions, and thyroid nodules. Ultrasound might also be useful for confirming the complete disappearance of superficial lesions.
(5) **Endoscopy and laparoscopy**

The utilization of endoscopy and laparoscopy for objective tumor evaluation has not yet been fully and widely validated. Their uses in this specific context require sophisticated equipment and a high level of expertise that may only be available in some centers. Therefore, the utilization of such techniques for objective tumor response should be restricted to validation purposes in specialized centers. However, such techniques can be useful in confirming complete pathological response when biopsies are obtained.

(6) **Tumor markers**

Tumor markers alone cannot be used to assess response. If markers are initially above the upper normal limit, they must normalize for a patient to be considered in complete clinical response when all lesions have disappeared. At present, studies are conducted to establish special criteria for the standardized use of PSA and CA-125 in clinical trials.

(7) **Cytology and histology**

Cytology and histology can be used to differentiate between CR (complete response) and PR (partial response) in rare cases such as germ cell tumors. In addition, cytology is needed in determining whether the overall evaluation should be PD (progressive disease) in the case that a measurable lesion was reduced or unchanged by treatment and appearance or increase of coelomic fluid was observed.

4. **Objective response evaluation**

(1) In objective response evaluation, tumors in all organs should be evaluated as a whole using the sum of their longest diameters, as contrasted to the evaluation based on the degree of response in each organ. If the protocol uses objective response as the primary end point, the study should include only the patients having measurable tumors at baseline.

(2) **Documentation of target and non-target lesions at baseline**

[Target lesions]: All measurable lesions up to a maximum of five lesions per organ and 10 lesions in total should be identified as target lesions. Target lesions should be selected based on their longest diameter and their suitability for accurate repeated measurements. Objective response should be evaluated based on the sum of the longest diameter for all target lesions.

[Non-target lesions]: All other lesions should be identified as non-target lesions. Measurements of these lesions are not required, but the presence or absence of each should be noted throughout follow-up.

(3) **Criteria for objective response**

(i) **Evaluation of target lesions**

Complete response (CR): Disappearance of all target lesions.

Partial response (PR): At least a 30% decrease in the sum of the longest diameter of target lesions, taking as reference the baseline sum of the longest diameter.

Progressive disease (PD): At least a 20% increase in the sum of the longest diameter of target lesions, taking as reference the smallest sum of the longest diameter recorded.

Stable disease (SD): Neither CR nor PR.

(ii) **Evaluation of non-target lesions**

Complete response (CR): Disappearance of all non-target lesions and normalization of tumor marker level.

Incomplete response (IR)/stable disease (SD): Persistence of one or more non-target

<table>
<thead>
<tr>
<th>Table 1 Overall Response Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target Lesions</td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>CR</td>
</tr>
<tr>
<td>CR</td>
</tr>
<tr>
<td>PR</td>
</tr>
<tr>
<td>SD</td>
</tr>
<tr>
<td>PD</td>
</tr>
<tr>
<td>Any</td>
</tr>
<tr>
<td>Any</td>
</tr>
</tbody>
</table>


lesion(s) or and maintenance of tumor marker level above the normal limits.

Progressive disease (PD): Unequivocal progression of non-target lesions.

(iii) Overall objective response

Overall response is evaluated based on the effect on target lesions and non-target lesions and the presence or absence of new lesions (Table 1).

(iv) Evaluation of best overall response

The best overall response is the best response recorded from the start of the treatment until disease progression. In general, the patient’s best response assignment will depend on the overall response and its duration.

(4) Duration of objective response

(i) Duration of CR: The duration of CR is measured from the time measurement criteria are met for CR until the date of recurrence (at least 4 weeks as defined in the study protocol).

(ii) Duration of PR: The duration of PR is measured from the time measurement criteria are met for PR until the date of disease progression (at least 4 weeks as defined in the study protocol).

(iii) Duration of SD: SD is measured from the start of the treatment until the criteria for disease progression are met (usually 6 to 8 weeks as defined in the study protocol). The clinical relevance of the duration of SD varies for different tumor types and grades. Therefore, it is highly recommended that the protocol specify the minimal time interval required between two measurements for determination of SD.

Features of New Toxicity Criteria (NCI-CTC Version 2.0)

1. Increased number of toxicities

Toxicities are classified into pathophysiological categories or anatomical categories. The 24 categories (Table 2) and 279 toxicities are arranged in alphabetical order.

2. Important changes in Version 2.0

(1) Infection has been subdivided into 4 categories according to absolute neutrophil count (ANC): “febrile neutropenia” (fever of unknown origin without clinically or microbiologically documented infection) (ANC<1,000/mm³, fever ≥38.5°C), “infection (documented clinically or microbiologically) with grade 3 or 4 neutropenia” (ANC<1,000/mm³), “infection without neutropenia,” and a new criterion “catheter-related infection” has been added.

(2) Associated toxicities should also be graded. Example: If toxicity “hemorrhage/bleeding with grade 3 or 4 thrombocytopenia” is noted, the criterion directs: “Also grade the site or type of hemorrhage/bleeding” and “Also consider Platelets, Hemoglobin, Transfusion-platelet, Transfusion-pRBCs.”

(3) The grade-4 criterion for “platelets” has been revised from <25,000/mm³ to <10,000/mm³.

(4) Toxicity Module and Infection Module have

---

Table 2  Toxicity Categories of NCI-CTC (Version 2.0)

<table>
<thead>
<tr>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALLERGY/IMMUNOLOGY</td>
</tr>
<tr>
<td>AUDITORY/HEARING</td>
</tr>
<tr>
<td>BLOOD/BONE MARROW</td>
</tr>
<tr>
<td>CARDIOVASCULAR (ARRHYTHMIA)</td>
</tr>
<tr>
<td>CARDIOVASCULAR (GENERAL)</td>
</tr>
<tr>
<td>COAGULATION</td>
</tr>
<tr>
<td>CONSTITUTIONAL SYMPTOMS</td>
</tr>
<tr>
<td>DERMATOLOGIC/SKIN</td>
</tr>
<tr>
<td>ENDOCRINE</td>
</tr>
<tr>
<td>GASTROINTESTINAL</td>
</tr>
<tr>
<td>HEMORRHAGE</td>
</tr>
<tr>
<td>HEPATIC</td>
</tr>
<tr>
<td>INFECTION/FEBRILE NEUTROPENIA</td>
</tr>
<tr>
<td>LYMPHATICS</td>
</tr>
<tr>
<td>METABOLIC/LABORATORY</td>
</tr>
<tr>
<td>MUSCULOSKELETAL</td>
</tr>
<tr>
<td>NEUROLOGY</td>
</tr>
<tr>
<td>OCULAR/VISUAL</td>
</tr>
<tr>
<td>PAIN</td>
</tr>
<tr>
<td>PULMONARY</td>
</tr>
<tr>
<td>RENAL/GENITOURINARY</td>
</tr>
<tr>
<td>SECONDARY MALIGNANCY</td>
</tr>
<tr>
<td>SEXUAL/REPRODUCTIVE FUNCTION</td>
</tr>
<tr>
<td>SYNDROMES</td>
</tr>
</tbody>
</table>

---

JMAJ, December 2003—Vol. 46, No. 12 557
been included to be implemented when more detailed information is considered pertinent.

(5) Options covering special therapies (hematopoietic stem cell transplant and the RTOG/EORTC Late Radiation Morbidity Scoring Scheme) have been added.

(6) Toxicity from radiotherapy occurring within 90 days after the start of radiation therapy is graded according to CTC, while that occurring greater than 90 days after radiation therapy is graded according to the RTOG/EORTC Late Radiation Morbidity Scoring Scheme.

Conclusion

Objective response evaluation in RECIST has been simplified from the method using bidimensional measurement (product of 2 diameters) in WHO Standards to the summation of unidimensional measurement (largest diameter). RECIST has already been used in some clinical trials on lung cancer in Japan. In time, response evaluation criteria will be totally based on RECIST.

NCI-CTC Version 2.0 includes toxicities and categories that have been increased from 49 toxicities in 18 categories to 279 toxicities in 24 categories, and the ability for appropriate evaluation of toxicity has been improved. The Japanese translation by JCOG (2nd edition)\textsuperscript{5} is now widely used in Japan.

REFERENCES