Original article

Annals of Oncology 15: 257–260, 2004
DOI: 10.1093/annonc/mdh059

Modified RECIST criteria for assessment of response in malignant pleural mesothelioma

M. J. Byrne[1]* & A. K. Nowak[1,2]

[1]*Department of Medical Oncology, Sir Charles Gairdner Hospital, Nedlands, WA 6009;* [2]*Department of Medicine, University of Western Australia, Nedlands, WA 6009, Australia*

Received 29 May 2003; revised 5 August 2003; accepted 30 September 2003

Background: The growth pattern of malignant pleural mesothelioma makes the use of RECIST (response evaluation criteria in solid tumours) response criteria difficult. We have developed and validated Modified RECIST criteria adapted to the growth pattern of malignant pleural mesothelioma.

Patients and methods: We evaluated 73 patients from two clinical trials of cisplatin/gemcitabine chemotherapy in malignant pleural mesothelioma. Tumour thickness perpendicular to the chest wall or mediastinum was measured in two positions at three separate levels on thoracic CT scans. The sum of the six measurements defined a pleural unidimensional measure. Bidimensionally measureable lesions were measured unidimensionally as for RECIST. All measurements were added to obtain the total tumour measurement. A reduction of at least 30% on two occasions 4 weeks apart defined a partial response; an increase of 20% over the nadir measurement, progressive disease. The validity of the modified criteria was gauged by evaluating survival and pulmonary function.

Results: Response according to these criteria predicted for superior survival (15.1 versus 8.9 months; $P = 0.03$) and forced vital capacity (FVC) increase during treatment ($P < 0.0001$). A significant correlation between change in linear tumour measurement and FVC was seen ($R = 0.63$; $P = 0.0001$).

Conclusion: These Modified RECIST criteria for tumour response correlate with survival and lung function and can be used to measure outcome in pleural mesothelioma.

Key words: chemotherapy, clinical trials, mesothelioma, RECIST, response criteria, validation

Introduction

The ability to measure reproducibly tumour response to treatment is vital in the development of new drugs and therapeutic combinations, particularly for the conducting of phase II studies. Conventional response criteria have always been difficult to apply to malignant mesothelioma due to its unique pattern of growth.

Malignant mesothelioma most commonly grows as a 'rind' around the pleural surface, and on computed tomography (CT) scan may not produce spherical lesions with bidimensionally measurable diameters. The WHO criteria [1] are poorly suited to evaluating response in mesothelioma, as they were developed principally to assess bidimensionally measurable disease. When used for unidimensional measurement, these criteria require a 50% decrease in the sum of unidimensional measurements to define a partial response (PR). This equates to a 75% decrease in the sum of the products of perpendicular diameters rather than the 50% decrease required to define response in bidimensionally measurable lesions.

**Correspondence to:* Dr M. J. Byrne, Department of Medical Oncology, Sir Charles Gairdner Hospital, Nedlands, WA 6009, Australia.
Tel: +61-8-93463841; Fax: +61-8-93463390;
E-mail: mjbyrne@cyllene.uwa.edu.au

The recently developed RECIST (response evaluation criteria in solid tumours) criteria [2] are more suited to tumour assessment in mesothelioma, as they specify the use of unidimensional measurements, with PR defined as a 30% decrease in the sum of the longest diameter for all target lesions. However, the selection of measurement sites in mesothelioma is difficult, and without further definition of the method of measurement, the RECIST criteria could be applied differently by different investigators (Figure 1A). Early experience suggests that modification of the criteria may be required in the special case of mesothelioma [3].

We have developed a modification of the RECIST criteria specifically to address the difficulties of measurement inherent in assessing changes in tumour bulk in pleural mesothelioma. We have previously performed two phase II clinical trials of the use of cisplatin and gemcitabine in patients with measurable pleural mesothelioma; a single centre study [4] and a subsequent confirmatory multicentre study [5]. In these studies, response rates of 47.6% [4] and 33% [5] were seen. Response criteria that incorporated both unidimensional and bidimensional measurements had been developed and used to assess response in these two trials ('original response criteria'). Using Modified RECIST criteria, we have now reassessed the response to treatment for a total of 73 patients who entered these two studies. To validate these new

criteria we have related response to survival and to serial changes in pulmonary function on treatment.

Patients and methods

Patient population

Data on the 73 patients entered in the two previous clinical trials of cisplatin and gemcitabine in malignant pleural mesothelioma were obtained. All patients had measurable disease, defined as pleural tumour thickness of at least 1.5 cm on spiral CT scan, and histologically or cytologically confirmed mesothelioma. CT scans had been performed in all patients prior to the first cycle of chemotherapy, and then before the second, fourth and sixth cycles. Further CT scans were performed after the final cycle of chemotherapy, and then twice a month until disease progression. Forced vital capacity (FVC) had been measured prior to study entry and on day 1 of each chemotherapy cycle in 52 patients entered on the second study.

The original response criteria used in both trials were as follows. Tumour measurements were performed on transverse cuts on thoracic CT scans at three separate anatomically reproducible levels on the study entry CT scan and at the same levels on subsequent scans. Where possible, bidimensional lesions were measured. If there were no bidimensionally measurable lesions, unidimensional measurements of pleural tumour thickness were performed. Bidimensionally measurable lesions were measured using the longest dimension and the length perpendicular to the longest measurement. For unidimensionally measurable lesions, thickness of pleural tumour was measured at two separate sites on each of the three levels and the six measurements summated to produce a total measurement. Palpable masses were measured clinically on day 1 of each cycle as for bidimensionally measurable lesions. A pleural effusion was not considered a measurable lesion.

Tumour response was defined as: (i) complete response (CR): disappearance of all known disease, determined by two observations not less than 4 weeks apart; (ii) PR: a ≥50% decrease in the sum of the products of perpendicular diameters of bidimensionally measured lesions on two occasions not less than 4 weeks apart, or a ≥30% decrease in the sum of linear tumour measurements on two observations not less than 4 weeks apart; (iii) no change: a decrease in bidimensional tumour area of <50% or an increase of <25%, or a decrease in the sum of unidimensional measurements of <30% or an increase of <25%, provided no new lesions have appeared; (iv) progressive disease: a ≥25% increase in the size of the tumour being measured (unidimensional or bidimensional) or the appearance of new lesions.

Where response at sites measured bidimensionally differed from that at sites measured unidimensionally, the overall patient response was assessed by an audit group. The sites with dominant tumour bulk were favoured.

The protocols for each study were approved by the Committee for Human Rights of the University of Western Australia and the Sir Charles Gairdner Hospital Clinical Drug Trials Committee. Written informed consent was obtained from each patient before entry.

Modified RECIST criteria

Modified RECIST criteria were developed. The major problems in applying the RECIST criteria to malignant pleural mesothelioma are in the interpretation of the meaning and placement of the 'longest unidimensional diameter' of the target tumour mass to be measured. The longest diameter of a tumour mass is frequently that which follows the inner curve of the chest wall. Defining the limits of such a diameter is often problematical. When the tumour regresses with treatment the line may cross an area outside the tumour margin because of the curve of the chest wall. This may produce difficulty with reproducibility of measurement. Furthermore, the longest tumour diameter may be between two fixed structures, such as the thoracic vertebrae and the carina, and measurement in these areas may not fully reflect tumour response.

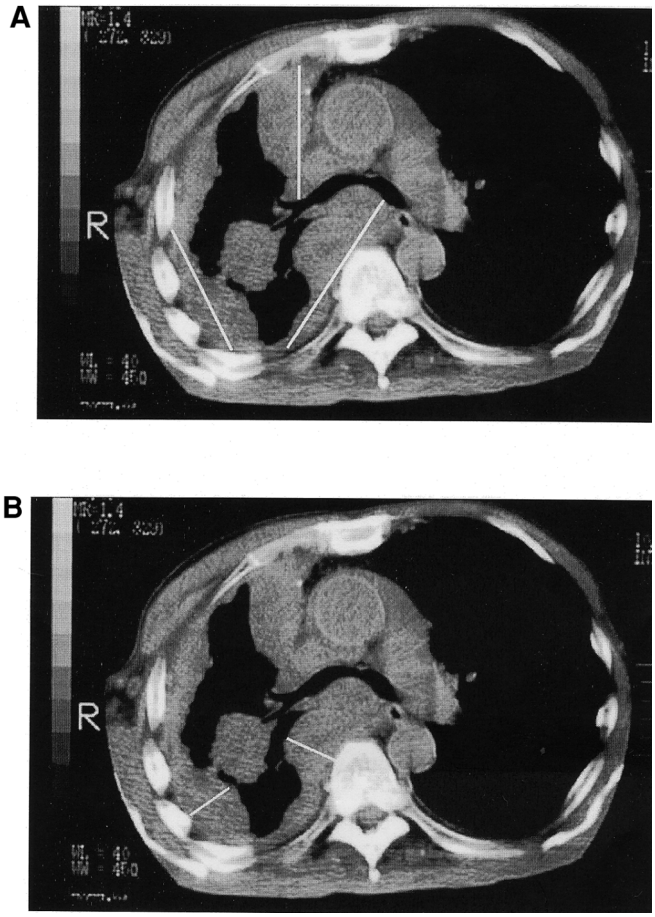

Figure 1. Example of measurement of a single computed tomography scan slice. (**A**) Lines represent possible interpretations of 'longest tumour diameter' according to current RECIST criteria. (**B**) Lines represent suggested measurement sites perpendicular to fixed structures, chest wall and vertebral column, according to Modified RECIST criteria.

The Modified RECIST criteria we have developed were as follows. Tumour thickness perpendicular to the chest wall or mediastinum was measured in two positions at three separate levels on transverse cuts of CT scan (Figure 1B). The sum of the six measurements defined a pleural unidimensional measure. Transverse cuts at least 1 cm apart and related to anatomical landmarks in the thorax were chosen to allow reproducible assessment at later time points. If measureable tumour was present, transverse cuts in the upper thorax, above the level of division of the main bronchi were preferred. At reassessment, pleural thickness was measured at the same position at the same level and by the same observer. This was not necessarily the greatest tumour thickness at that level. Nodal, subcutaneous and other bidimensionally measurable lesions were measured unidimensionally as per the RECIST criteria. Unidimensional measurements were added to obtain the total tumour measurement.

CR was defined as the disappearance of all target lesions with no evidence of tumour elsewhere, and PR was defined as at least a 30% reduction in the total tumour measurement. A confirmed response required a repeat observation on two occasions 4 weeks apart. Progressive disease (PD) was defined as an increase of at least 20% in the total tumour measurement over the nadir measurement, or the appearance of one or more new lesions. Patients with stable disease (SD) were those who fulfilled the criteria for neither PR nor PD.

Table 1. Comparison of response at individual time point observations between original response criteria and Modified RECIST criteria

Original response criteria	Modified RECIST criteria				
	CR	PR	SD	PD	Total
CR	0	0	0	0	0
PR	0	72	5	4	81
SD	0	11	93	1	105
PD	0	1	2	47	50
Total	0	84	100	52	236

RECIST, response evaluation criteria in solid tumours; CR, complete response; PR, partial response; SD, stable disease; PD, progressive disease.

Table 2. Comparison of response rates for original response criteria versus Modified RECIST criteria

	CR	PR	SD	PD	Overall response (%)
Original response criteria	0	27	40	6	37
Modified RECIST criteria	0	27	40	6	37

RECIST, response evaluation criteria in solid tumours; CR, complete response; PR, partial response; SD, stable disease; PD, progressive disease.

Validation of modified criteria

The response status of all 73 patients was re-assessed at each trial time point according to these Modified RECIST criteria. Patients were assigned to one of two groups: responding patients ('responders') and patients with SD or PD ('non-responders'). Overall survival from the start of treatment and serial changes in FVC were analysed for these two groups as a surrogate measure of patient benefit.

Results

Tumour measurements from a total of 236 CT scans from 73 patients were reassessed for response status at each time point in the trials according to the modified criteria. In general there was a close correlation between the original response criteria and the Modified RECIST. Of 81 time points originally showing PR, 72 remained classified as PR, five became SD and four became PD. Of 105 time points originally showing SD, 93 remained as SD, 11 became PR and one became PD. Of 50 time points originally showing PD, 47 remained PD, two became SD and one became PR (Table 1). The overall confirmed response rate as assessed by the two systems did not differ, as two patients classified as SD became PR and two patients classified as PR became SD (Table 2).

Median survival was plotted as a Kaplan–Meier curve for the responding and non-responding patients (Figure 2). There was a statistically significant difference in survival between the two patient groups (log rank test $P = 0.03$). Median survival was 15.1 months for responding patients and 8.9 months for non-responding patients.

FVC was plotted as a percentage of starting FVC for the responding and non-responding patients (Figure 3). FVC improved

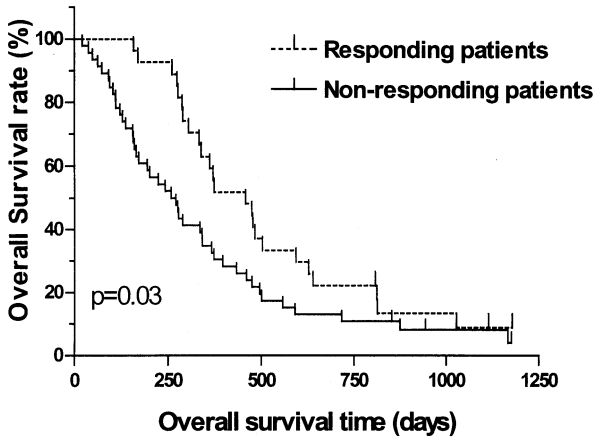

Figure 2. Kaplan–Meier survival curve of overall survival (days) from start of treatment for 73 patients treated with cisplatin and gemcitabine. Responding (dashed line) versus non-responding (solid line) patients as per Modified RECIST criteria. *P* value represents results of log rank test.

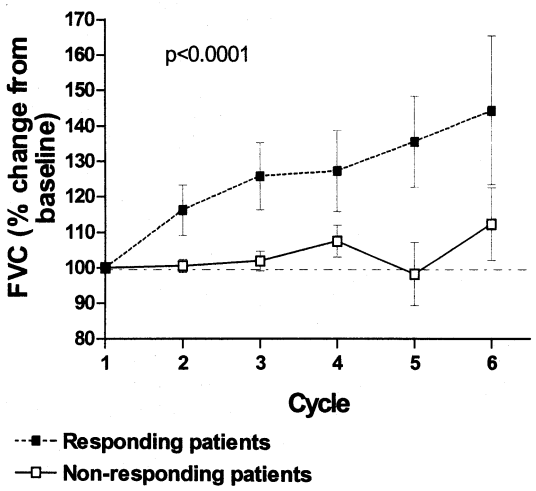

Figure 3. Forced vital capacity (FVC) as percentage of individual baseline FVC for all 73 patients treated with cisplatin and gemcitabine. Responding (filled squares) versus non-responding (open squares) patients as per Modified RECIST criteria. *P* value represents results ANOVA.

significantly over the course of treatment for responding patients as compared with non-responding patients ($P = 0.0001$).

The change in FVC against change in linear tumour measurement was plotted for each patient (Figure 4); a significant correlation was seen ($R = 0.63$, $P = 0.001$).

Discussion

The response criteria used in our previous two trials of chemotherapy in mesothelioma incorporated bidimensional measurements, as the WHO response criteria were widely used when the first of these trials began accrual. These response criteria were not altered for the second trial, which sought to broaden the applicability of our findings in the first single-centre study into a

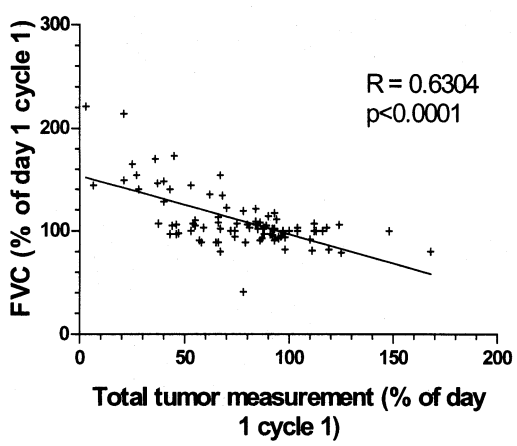

Figure 4. Forced vital capacity (FVC) versus total tumour measurement standardised as percentage of individual baseline on day 1 of cycle 1. $R = 0.6304$, the result of linear regression analysis; $P < 0.0001$, significance of deviation from 0 of the line of regression.

multicentre setting, adding prospective evaluation of quality of life and lung function end points. In order to make the response rates of the two trials directly comparable, it was important that the measurement criteria used in the two studies were the same. Subsequently, the RECIST criteria have been developed and become widely accepted and used in clinical trials. The application of RECIST criteria could, however, be variably interpreted by different investigators in mesothelioma, and this may lead to unsatisfactory results [3]. We have developed Modified RECIST criteria that are specifically designed to address the unique growth pattern of pleural malignant mesothelioma. Use of these modified response criteria did not materially alter the response rates in our two previous trials. However, the modified criteria avoided difficult and ambiguous situations that arose in response interpretation in the two previous trials; for example, how to evaluate response when discordance between unidimensional and bidimensional lesions occurred.

Whilst tumour response and progression directly reflect changes in tumour bulk, they are most clinically useful when they relate closely to other measures of a patient's condition. Response is a surrogate for patient benefit in the evaluation of new drugs and combinations. Patient benefit in pleural mesothelioma may include an improvement in survival or lung function, and improvement in symptom control or quality of life. Thus, it is important that any valid measurement of response should reflect changes in these parameters. We have demonstrated that these Modified RECIST criteria successfully distinguish between responders and non-responders for the parameters of survival and change in FVC, thus demonstrating their validity.

Further evaluation of these modified criteria should be performed before they can be incorporated routinely into future clinical trials. The development of an automated measurement format may enhance the speed and reproducibility of the measurements [6] and go some way to overcoming the potential problem of inter-observer variability. We have avoided this issue by using the same observer or an audit group to undertake the measurements. Tests of inter-observer variability are important, however, as has been demonstrated in the assessment of response in lung cancer [7]. To confirm the practicality of the criteria they should be applied to a group of clinicians who lack extensive experience in the measurement of mesothelioma in clinical trials.

Acknowledgements

A.K.N. was the recipient of an Eva K. A. Nelson scholarship.

References

1. Miller AB, Hoogstraten B, Staquet M, Winkler A. Reporting results of cancer treatment. Cancer 1981; 47: 207–214.
2. Therasse P, Arbuck SG, Eisenhauer EA et al. New guidelines to evaluate the response to treatment in solid tumors. J Natl Cancer Inst 2000; 92: 205–216.
3. van Klaveren RJ, Aerts JGJV, de Bruin HG et al. Inadequacy of the RECIST criteria for the evaluation of response in patients with malignant pleural mesothelioma (MPM). Proc Am Soc Clin Oncol 2002; 21: 310a.
4. Byrne MJ, Davidson JA, Musk AW et al. Cisplatin and gemcitabine treatment for malignant mesothelioma: a phase II study. J Clin Oncol 1999; 17: 25–30.
5. Nowak AK, Byrne MJ, Williamson R et al. A multicentre phase II study of cisplatin and gemcitabine for malignant mesothelioma. Br J Cancer 2002; 87: 491–496.
6. Armato SG, Oxnard GR, Macmahon H et al. A computer interface for the semi-automated measurement of mesothelioma on CT scans. Proc Am Soc Clin Oncol 2003; 22: 668.
7. Erasmus JJ, Gladish GW, Broemeling L 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: 2574–2582.