available at www.sciencedirect.com

SCIENCE DIRECT®

journal homepage: www.ejconline.com

Review

RECIST revisited: A review of validation studies on tumour assessment

P. Therasse, E.A. Eisenhauer, J. Verweij*

EORTC Data center, Avenue E. Mounier, 83 Box 11, 1200 Brussels, Belgium

ARTICLE INFO

Article history:
Received 22 January 2006
Accepted 24 January 2006
Available online 17 April 2006

Keywords:
RECIST
WHO
Literature review
Clinical trials

ABSTRACT

The response evaluation criteria in solid tumours (RECIST) was developed in the late 1990s to replace the WHO criteria for response evaluation. The new criteria included important changes such as unidimensional tumour measurement, selection of target lesions with a minimum size, details concerning imaging modalities and a new threshold for assignment of objective progression.

RECIST was published in February 2000 and very quickly came into operation first in clinical trials performed under the auspices of EORTC, US NCI or NCI Canada Clinical Trials Group but was adopted quickly thereafter by the entire cancer clinical research community. As several key features of RECIST were based on analysis of retrospective clinical data, it was felt important to carefully monitor the implementation of the guidelines and stimulate prospective validation studies. This paper reviews the literature that has been published on RECIST from 2000 up to November 2005. In total 60 papers and ASCO, abstracts directly refer to research studies or reviews related to RECIST and its implementation. Amongst the 60 references identified for this review, 11 papers refer to validation studies (seven prospective and four retrospective), six papers refer to the comparison of unidimensional measurements versus bi or tri-dimensional measurements, 12 papers address issues raised with the implementation of RECIST in Mesothelioma and Gastro-Intestinal Stromal Tumours and four papers report on an adaptation of RECIST for specific tumour types.

In general, RECIST has been well received by the scientific community and most validation studies fully support the implementation of the new criteria. As expected, however, some issues have been identified. In keeping with the mathematical differences in definition of progression, RECIST delays the identification of progression as compared to WHO criteria in some instances. RECIST criteria are not easily applicable in some types of trials such as those in paediatric tumours and in mesothelioma. Furthermore, anatomical changes in the tumour as described by RECIST may be detected later than functional changes in some circumstances, as for example in Gastro-Intestinal Stromal Tumours treated with Imatinib. However, there is no other universal method of tumour assessment as yet and functional imaging methods have not been validated and will not be widely available for some time. The findings of this review, together with experience acquired thus far and the results of some ongoing research projects, have paved the way for RECIST 2.0 to be hopefully announced later this year.

* *Corresponding author:* Tel.: +32 2 774 16 14; fax: +32 2 772 61 97.
E-mail address: patrick.therasse@eortc.be (P. Therasse).
0959-8049/$ - see front matter
doi:10.1016/j.ejca.2006.01.026

1. Introduction

Response evaluation criteria in solid tumours (RECIST) was introduced by a small international working group in February 2000 to facilitate, improve and standardize the evaluation and the reporting of objective tumour outcomes in early clinical trials investigating new anti-cancer agents.[1] In comparison to earlier response assessment systems, the new criteria gave much more detailed recommendations on how to assess tumour lesions, how to report responses, and also took into account recent developments in medical imaging techniques. RECIST uses a unidimensional measure (the longest diameter) to quantify measurable tumour lesions as opposed to the bidimensional product (longest diameter multiplied by its perpendicular), which was commonly employed by earlier iterations of response criteria.[2–4] Building on the work of others,[3,5] RECIST defines measurable lesions as those with a minimum size depending on the method of investigation. Following a principle already implemented in the SWOG response criteria,[3] the threshold for defining objective progression was arbitrarily increased as compared to the WHO criteria, i.e., the increase in measurable overall tumour burden required for progression was greater in RECIST (20% in one dimension being approximately equivalent to a 44% increase in bidimensional product) than in the WHO criteria (25% increase in product).

Following the publication of RECIST, standard case report forms (CRFs) and protocol sections were created by the working group and made available on the web. A special email address was created to receive and answer questions related to the implementation of the criteria. A website was created to host the Questions and Answers to facilitate the implementation of the criteria (www.eortc.be\recist). Although the last comment on the website was posted in 2003, the RECIST working group continues (weekly) to answer questions and provide support for the interpretation of the criteria in specific situations.

After the publication of RECIST, some investigators raised concerns about the interest, the pertinence and the applicability of the new criteria. The main purpose of this paper is to review the work performed and published by other colleagues on the usefulness of the criteria in general and their validation in specific tumour types when available.

2. Review methodology

The search strategy was simple and made through PUBMED using the word RECIST as keyword to identify titles and abstracts published between February 2000 and November 2005. This search strategy identified 99 referenced papers. Only those manuscripts reporting on original work focused on the methodology of response evaluation and RECIST were retained for detailed review. Also excluded were editorial comments and non-English literature. Ultimately 43 papers satisfied these criteria. A second search was undertaken of abstracts published in the American Society of Clinical Oncology (ASCO) annual conference proceedings between 2001 and 2005. This identified a further nine abstracts (and related data in oral presentations or posters) that had not yet been followed by a full paper. Finally, examination of the reference lists in the 43 full papers yielded another eight additional papers which met the review criteria. Thus in total, 60 studies (51 papers and nine ASCO abstracts) were identified for inclusion in this review.

3. Results

The studies included focused either on general principles relating to the implementation of RECIST (or tumour evaluation) or on a prospective or retrospective attempt to validate the utility of RECIST in certain tumour types. Accordingly, the results of this review have been divided into general and tumour specific considerations.

3.1. *General considerations*

One of the first papers to refer to RECIST was a commentary of Padhani and Husband.[6] The authors outlined the problems inherent to the morphological assessments of tumours independently of the number of dimensions being measured and briefly explored the development of functional imaging as a tool for the future. However, their conclusion was crystal clear: “current criteria should remain unchallenged until better functional parameters emerge”. One year later the same first author[7] analysed RECIST and its impact on radiology departments highlighting the possibility that the implementation of RECIST could translate into increased workload. The paper concluded that, while the issue of workload required careful monitoring, this factor alone should not be an argument to be less stringent in response assessment in the performance of clinical trials. Institutions that could not provide this service should be considered incapable of performing studies where response assessment is crucial. In 2004, the International Cancer Imaging Society (ICIS) published a consensus statement about the evaluation of the response to treatment of solid tumours,[8] including a number of issues related to the implementation of RECIST (Table 1). Another paper[9] published almost simultaneously but in another journal identified very similar issues. It is interesting to note that on one hand these authors cite concern about the potential increase in workload created with the application of RECIST (specifically the requirement to measure up to 10 lesions if multiple measurable lesions are identified), while on the other hand advise consideration for the use of 3D measurements. Three dimensional measurements to date have not been shown to be more useful than 1D measurements (for the purpose of response evaluation), but is certainly much more complex and time consuming.

The general concordance between RECIST and WHO criteria was tested retrospectively in a cohort of 130 patients with different tumour types and entered into different protocols.[10] In line with the larger increase in lesion size required for definition of progressive disease (PD) found in RECIST , it was shown that about 1/3 of patients normally identified as PD with WHO criteria would still be classified as having stable disease (SD) with RECIST. The authors also used this dataset to create multiple simulations to artificially change tumour shape to demonstrate that increasing the irregularity of lesions may decrease the concordance rate of partial response (PR) and SD categories between the two methods.

Table 1 – Main concerns expressed by the International Cancer Imaging Society concerning RECIST

	Concerns
General	RECIST is a first step in the good direction but requires well trained radiologists who should also be involved in the planning of trials
WHO vs. RECIST	Several studies have shown a good concordance between RECIST and WHO for response but less good concordance for time to progression. This should be taken into account for planning of future trials
One dimension vs. three dimensions	Three-dimensional imaging is available in many centers and should be considered in the guideline
Specific issues	• Lymph-node should be measured in the short axis • Changes in tumour consistency (calcification, necrosis...) should also be reported for accurate evaluation of tumours • Cystic lesion should not be systematically excluded • Why 10 target lesions? No scientific rationale • Appendix concerning imaging technology should be updated • MRI could be used to measure bone disease (breast, prostate) • More information should be given on the use of contrast for imaging

In an analysis of 32 North Central Cancer Treatment Group (NCCTG) trials including 2374 patients,[11] it was suggested that two lesions were sufficient to provide a reliable tumour assessment. However only 23% of the patients studied had more than two lesions at baseline. Schwartz and colleagues[12] used the data of 36 patients with multiple target lesions to simulate hundreds of possible groupings between target lesions and suggested that when six lesions were measured bidimensionally and four lesions were measured unidimensionally, the average variance was decreased by 90%. The same authors subsequently developed a mathematical model[13] which could in theory enable physicians to calculate the optimal number of lesions to follow to decrease the variance to an acceptable level, for patients with a large number of lesions at baseline. Although this concept is certainly interesting, it could increase the workload for response assessment so it would be strengthened if it underwent further testing on a large database generated on prospective studies.

Perez-Gracia and colleagues[14] analysed nine phase II trials (same drug but different tumour types) including 416 patients amongst which 97 responses were first recorded and 81 were later confirmed. Most unconfirmed responses were due to patients lost to follow-up. Because of the high correlation coefficient between the rate of confirmed and non-confirmed responses, the authors suggest that confirmation of response may not be necessary and should be studied in a larger setting. The relevance of response confirmation is part of an ongoing debate. The need for confirmation of response should be discussed as an approach for further simplifying the RECIST.

Three papers from the same group of authors addressed the problems of the implementation of RECIST in paediatric oncology. Two papers[15,16] focussed on technical problems while the third paper[17] provided a short illustrative description of 10 cases and the issues raised by attempting to measure the tumour response according to RECIST. A summary of the most important points raised is presented in Table 2. There are very few studies with a focus on the methodology of response assessment in paediatric solid tumours and the problems reported by the authors seem to relate more to the application of RECIST in clinical practice than to the use of RECIST for screening potential new anti-cancer agents in clinical trials. It is not uncommon that individual patients (adult and paediatric) in oncology practice cannot have response assessed according to RECIST if they do not have measurable disease: these cases would not normally be eligible for inclusion in phase II trials where response is the endpoint.

In summary, the major issues identified in these reports were related to the need for measurement of 10 lesions, the fact that some patients with progression by WHO were considered to have stable disease by RECIST, proposals for different approaches to lesion measurement, and the minimal impact in terms of overall response rate of the requirement for response confirmation.

3.2. *Tumour specific considerations*

3.2.1. *Lung cancer*

Three papers compared RECIST and WHO criteria for the assessment of response in Non-Small Cell Lung Cancer.[18–20] The results of the three studies are summarized in Table 3. They report a good correlation between unidimensional measurements and sum of bidimensional products in keeping with the data supplied in the RECIST paper. The application of RECIST translated into an ineligibility rate of 5% because

Table 2 – Main concerns raised by paediatricians regarding the implementation of RECIST in paediatric studies

	Concerns
General	• Problems related to disseminated disease with diffuse infiltration • The minimum size of target lesions should be smaller than 10 mm with current available techniques (multislice CT)
Imaging	• The need for repetitive exposure to radiation burden when multiple CT should be performed, hence ultrasonography avoids unnecessary radiation exposure and ensures much better compliance for children • Bone lesions should be acceptable as they can be evaluated by MRI • All radiological plans should be considered to measure tumour lesions and not the axial plane only • Functional imaging should also be considered as a possible modality for tumour evaluation in some instances

of the requirement for a minimum size of the target lesion,[20] a factor on which WHO is silent. Three other papers[21–23] analyzed the intra- and inter-observer variability in tumour response evaluation using either RECIST or RECIST and WHO criteria but did not make a comparison in terms of response and progression rates between the two sets of criteria. In all three papers, the unidimensional approach reduced the inter-observer variability of the measurements and the misclassification of some patients. This is further supported by the observation that the inter-observer variability improves when minimum lesion size criteria are used.[20] Interestingly, the papers were almost unanimous in recommending RECIST for tumour evaluation in future trials in NSCLC.

3.2.2. *Mesothelioma*

Tumour evaluation in mesothelioma has always been a difficult problem, with the original WHO criteria being modified in 1997[24] to enable the use of unidimensional measurements for response assessments. Despite this adaptation, the modified WHO criteria still do not seem optimal given the frequent discordance between the evaluation of bidimensionally measured lesions and unidimensionally measured lesions.[25] One might have expected that RECIST, focusing only at unidimensional measurement, would be less confusing.[26] Yet, in two small studies (34 patients and four patients, respectively),[27,28] there was considerable discordance between the evaluation performed according to either RECIST or WHO criteria, both for objective response and progression. Byrne[25] proposed modified RECIST criteria for mesothelioma (using the longest perpendicular diameter to chest wall or mediastinum measured at two sites at three different levels on CT scan). He tested these modified RECIST criteria against the more complex WHO modified criteria retrospectively with data from two trials and found a good correlation between them. Responses reported with the modified criteria were associated with longer survival and improved lung function. Similar outcomes had been reported by others.[29] Further testing in reasonably sized prospective trials is encouraged for the definitive validation of the modified RECIST for mesothelioma.

3.2.3. *Breast and colorectal cancer*

Four studies[30–33] have compared the response rates observed with WHO criteria and RECIST in metastatic breast or colorectal cancer (Table 3). Interestingly, the study by Trillet-Lenoir[30] has investigated the original WHO criteria (as published) as well as a modified version (no progression declared when only one lesion is progressing and no more than five lesions measured per organ) which correlated very well with RECIST both for response and progression. In a second study,[31] although the overall response rates were identical between RECIST and WHO criteria, four patients were recorded with a worse response and four with a better response with RECIST. These were all patients with irregularly shaped tumours. Tran[34] compared the change in lung metastases measurements in 15 patients using WHO criteria (2D), RECIST (1D) and a 3D measurement system (3D). The author reported that the 1D and 3D measurements were concordant in 29 of 30 classifications, the 2D and 3D measurements were concordant in 23 of 30 classifications and 1D and 2D were concordant in 24 of 30 classifications. Despite the good concordance between 1D and 3D assessments, the level of agreement (measured with Kappa statistic) did not reach significance and the overall correlation between the various methods was considered fair to poor. The author acknowledged that many tumours were irregularly shaped and thus presented challenges in correctly calculating the 3D measurements which might in part explain the results. Kimura and colleagues[35] retrospectively assessed 50 breast cancer patients comparing RECIST with the standard criteria used by the Japanese Cancer Society for breast cancer

Table 3 – Prospective and retrospective studies comparing WHO criteria with RECIST

Author	Tumour type	P/R[a]	Sample size	RR (%) RECIST	RR (%) WHO	PD (%) RECIST	PD (%) WHO
Werner[18]	Locally advanced lung	P	22	87	87	NA	NA
Cortes[19]	Metastatic NSCLC[b]	R	164	52	52	26	26
Watanabe[20]	Metastatic NSCLC	R	120	19.3	20	13	17.5
Trillet Lenoir[30]	Metastatic Colorectal	P	91	25	20	41	43
Prasas[31]	Metastatic Breast	P	86	50	50	NA	NA
Choi[32]	Metastatic Colorectal	P	41	36	32	NA	NA
Muro[33]	Metastatic Esophageal	P	52	20	24	39	43
Therasse[42]	Metastatic STS[b]	P	49	6.1	4.1	32.6	34.7
Negrier[51]	Metastatic RCC[b]	P	61	14.7	11.4	44.3	57.3
Schwartz[49]	Metastatic RCC	R	53	5.6	3.7	11.3	17
Park[60]	Mix of tumours	R	79	30.4	31.6	30.4	38

a P, prospective; R, retrospective.
b STS, soft tissue sarcoma; RCC, renal cell cancer; NSCLC, non-small cell lung cancer.

evaluation. As the Japanese criteria are different from that of WHO, the meaning of the comparison with RECIST is not particularly clear. However, the author noted that 32% of patients would have been ineligible for studies had RECIST been used because of the minimum size for target lesions and the exclusion of patients with bone metastases only.

3.2.4. *Gastro-intestinal stromal tumour (GIST) and soft tissue sarcoma*

GIST is a relatively rare disease as compared to other cancers and until recently there has been no effective systemic therapy available. Imatinib, an oral tyrosine kinase inhibitor, has proven to be very effective to treat GIST. Early metabolic responses can be observed with ^{18}Fluorodeoxyglucose (FDG) positron emission tomography (PET), a technique that assesses metabolic activity in tissues. The FDG PET responses usually precede objective response by several weeks.[36] Many investigators have reported problems with the evaluation of GIST based on changes in tumour size.[37–41] GISTs can sometimes increase in size as a result of the metabolic response (intratumoural haemorrhage or mixoid degeneration) or intratumoural nodules can be reactivated showing an increased metabolic activity (translating early resistance to treatment) while the size of the tumour remains globally stable. FDG-PET has some limitations for evaluating tumour response to treatment, such as specificity, access, costs and quantitative measurements. Stroszczynski[40] proposes to associate MRI examinations with changes in tumour size and Choi[41] proposes a new set of criteria based both on changes in tumour size and changes in tumour density measured on CT images. The criteria proposed by Choi are currently being evaluated in a multicenter study.

Therasse[42] studied the outcome of 49 non-GIST soft tissue sarcoma patients treated for metastatic disease with ET-743. The response rates reported using RECIST and WHO based criteria were 6% and 4%, respectively, for partial response, 61% for no change (stable disease) using both methods and 33% and 35%, respectively, for progression. The outcome of the study would have been the same regardless of the criteria used. RECIST may correlate much less favourably with histological response in locally advanced high-grade soft tissue sarcoma treated with neoadjuvant doxorubicin and ifosfamide. In a total of 41 patients, 11 had good histologic response while only one patient had a response by RECIST.[43] Similar lack of correlation has also been observed with WHO criteria, so it may be that the neoadjuvant approach in soft tissue sarcomas is not suited for any size-related assessment of response.

3.2.5. *Prostate cancer*

Prostate cancer metastases are well known to be difficult to measure by imaging. Tombal[44] reported that only 11 (29%) of 38 consecutive metastatic hormone refractory prostate cancer patients had measurable disease according to RECIST. Twenty-five patients had focal metastatic bone lesions identified and potentially measurable on axial-skeleton MRI (see below). In another study,[45] two cohorts of patients with hormone refractory prostate cancer (HRPC) (31 patients) and hormone sensitive prostate cancer (HSPC) (124 patients), respectively, were analysed for eligibility to enter clinical trials in which RECIST criteria were to be used. Thirty-nine percent of HRPC patients and 51% of HSPC would have qualified on the basis of having at least one measurable lesion, while 13% and 44%, respectively, had bone-only disease (which is non-measurable). In this prostate study, most of the visceral measurable lesions were in lymph-nodes for which the evaluation by RECIST may also be difficult. Clearly, the evaluation of prostate cancer in screening phase II studies may require different tools, to determine outcome measures for clinical benefit defined according to stage of the disease and/or composite endpoints that include the various clinical dimensions of the disease (including measurable disease, QOL, PSA, Bone and pathology).[45]

3.2.6. *Brain tumours*

While in the past, brain tumours were thought to be difficult to assess for response, this currently seems less of an issue. Four separate studies[46–49] on a total of 204 patients showed a high concordance between 1D (RECIST), 2D (WHO) and 3D (volumetric) measurements in detecting responses both in childhood and adult brain tumours. For disease progression, the results were a little less uniform. One study suggested that progression was detected later with the 1D measurement.[46] Although three of these studies have as yet only been reported in abstract form, the data suggest that RECIST is a useful tool for brain tumour measurement.

3.2.7. *Renal cell carcinoma*

In two studies on metastatic renal cancer,[50,51] outcome according to RECIST and WHO criteria correlated extremely well (Table 3).

3.2.8. *Other issues*

The issue of functional imaging has already been mentioned. Gopinath[52] assessed functional volume variations against RECIST in 22 patients with neuroendocrine tumours of carcinoid type in the liver and treated with chemotherapy. He found out that tumour functional volume assessed by single-photon emission computed tomography (SPECT) predicted clinical outcome (as measured by a reduction in pain, flushing or abdominal symptoms) for 59% of patients and RECIST only for 36% of patients.

Another problem is the use of a primary tumour for response assessment, if the tumour is localized in a hollow organ. Indeed the nature of cancer growth for instance within the wall thickness of the oesophagus makes measurements based on RECIST difficult. Response to treatment is even more difficult to image because of post-treatment fibrous stenosis.[53] Tahara proposed to incorporate in RECIST a set of simple criteria based on endoscopic evaluation and histology to validate the complete response. The application of the proposed criteria correlates well with survival in a cohort of 139 patients reviewed retrospectively. RECIST was assessed in advanced gastric cancer by Yoshida[54]. He compared the response rate of 161 patients with measurable disease included in a phase III trial and concluded that response rates by the old and new criteria were nearly equal.

Following tumour markers is another possible approach to assessing change in tumour burden. Successive CA-125 measures were compared with sequential tumour measurements

as prognostic factor for survival in 131 patients receiving second-line chemotherapy for advanced ovarian cancer.[55] The study was retrospective and based on patients from a single institution treated with similar regimen in first and second-line. In this setting, CA-125 response was 2.6 times better than clinical response (assessed by RECIST) in predicting survival. However, in this example, the authors have examined both endpoints in relationship to their predictive value in survival in individual patients. This is a somewhat different application than use of the endpoint to screen new drugs for activity. The authors also confirm the difficulties in using radiographic techniques for assessing recurrent ovarian carcinoma.

3.2.9. *Metastatic site related considerations*

Bone metastases represent a frequent problem in breast and prostate cancer. Approximately 20–30% of these patients will present with bone metastases as the only metastatic site. Since RECIST consider bone metastases as non-measurable, several authors have tried to assess this problem.[7,9,35,44,45] While they confirmed the problem, they unfortunately could not provide alternatives. In contrast, Hamaoka[56] proposed a set of criteria requiring the combination of different imaging modalities (skeletal scintigraphy, plain radiography, computed tomography and/or magnetic resonance imaging) depending on the characteristics of the lesions. Tombal[44] investigated the possibility of using the axial-skeleton MRI (AS-MRI) for quantitative imaging of bone metastases of prostate cancer. He prospectively performed AS-MRI in a cohort of 30 patients with HRPCa and a positive Tc-99m bone scan before and six months after starting chemotherapy, and suggested that the proportion of patients eligible to enter trials based on RECIST increased from 28.9% to 65.7% with measurable lesions on AS-MRI. The feasibility of this approach should be tested prospectively taking into account technical issues, cost and time constraints.

Lymph-nodes and nodal masses can be manifestations of many cancers, but RECIST has yet only been reported compared to WHO in primary lymphomas. Sohaib[57] compared CT assessment in 1D, 2D and 3D for tumour response in 16 patients with either lymphoma or germ cell tumours. He concluded that whichever method is used, there is limited influence on the classification of treatment response. Assouline[58] pooled the data from three phase II lymphoma trials (115 patients) and compared RECIST with the International Working Criteria (IWC).[59] RECIST was slightly adapted to make the criteria relevant for response in lymphoma and the overall response rates were 42% and 46% for IWC and RECIST, respectively, with identical progression rates.

4. Discussion

RECIST has become the most frequently used response criteria for clinical trials investigating new treatments for solid tumours. The criteria are used to define response rate, progression rate and/or time to progression irrespective of the stage of development of new cancer therapeutics. Some features of the criteria have also been rapidly implemented in day to day practice of oncologists for standard patient care.

Overall, many authors agree that the development of RECIST with rigorous evaluation of many underlying aspects of response assessment has been very valuable. However, RECIST is not the universal panacea that one would like to have to precisely measure tumour response and progression in all possible situations and with all types of cancer therapeutics. Interestingly, the implementation of RECIST has also revealed a number of otherwise uncovered problems related to response evaluation in specific situations.[8,9,17,25,29,35,45] Although RECIST may not have provided an answer to all problems it has the merit of having stimulated discussion, and therefore improved awareness and harmonization.

The first objective of this review was to look into all prospective and retrospective studies attempting to validate RECIST against the WHO criteria. An overview of all studies directly comparing the two methods is presented in Table 3. In keeping with the retrospective data compiled in the RECIST manuscript, none of these studies found major differences in response rate between the two methods, while some found a slightly longer time to progression or lower progression rate for RECIST. The latter was expected and confirms that the changes in the definition of progression with RECIST translate into fewer patients being classified as having progression at a certain time as compared with WHO criteria. This should be taken into account when time to progression or progression free survival serve as a primary endpoint in non-randomized phase II trials where the primary hypothesis is constructed with reference to historical controls based on WHO criteria evaluations. This is of particular importance now that we move towards the development of drugs which may not induce rapid tumour shrinkage. Reference matrices for this purpose can be derived from recent large phase III, which have used RECIST for response and progression assessment. EORTC is currently developing such references for several tumour types.

Apart from directly comparing RECIST to WHO criteria, several groups have investigated the value of volumetric measurement (3D) as opposed to bidimensional (2D) (WHO criteria) or unidimensional (1D) (RECIST) approaches for assessing tumour response. There was a belief, particularly for lung and brain tumours, that volumetric assessment would be better. In addition, radiologists[8,9] and paediatricians[15–17] have raised concerns on the deficits of non-volumetric measurements. However, in all reported studies[18,34,46,47,49,57] the correlation between 1D and 3D dimension was quite good, with one possible exception for classifying progression in childhood cancers which would have been detected earlier with a 3D assessment.[46] However, the criteria for declaring progression were not proportionally identical. Apparently, the use of more laborious 3D measurements does not add great value to the much simpler 1D assessment of RECIST for the purpose of response determination in clinical trials.

RECIST has been found not optimal for assessment of response in mesothelioma and gastro-intestinal stromal tumours (GIST). Mesothelioma because of its specific growth characteristics deserves specific criteria, and some adaptations of RECIST as proposed by Byrne[25] seem an appropriate

compromise to keep criteria as simple as possible yet keeping a good correlation with clinical outcome. For GIST the problem is less well elucidated. The implementation of RECIST came in parallel with the clinical development of Imatinib in GIST and early treatment monitoring of this targeted therapy with FDG-PET identified a metabolic response much before any morphological response (tumour shrinkage) could be reported according to RECIST.[36–39] However, this is only a chronology issue and does not yet disqualify RECIST. And for many different reasons,[40,41] not least the limited availability of the tool, FDG-PET is at present not yet a universal method to assess tumour response in GIST. Yet, also in GIST, size changes might be opposite to clinical benefit observed, and early changes in tumour density without significant size changes may still harbour progression or response. This renders RECIST (as presently defined) not totally appropriate for use in GIST, and other criteria are currently being tested.[41]

The development of Imatinib shows that we have to be careful with the evaluation of new cancer therapeutics using standard tools. Evaluation solely based on morphological changes can be misleading and an accurate early determination of response may require functional and molecular techniques that assess metabolism, growth kinetics, angiogenesis growth factors, tumour cell markers and in vivo genetic alterations and gene expression. The US National Cancer Institute therefore has designed a very large research program to develop imaging in oncologic drug development.[61,62]

Other potential refinements and facilitations of the process of response evaluation[11–14] are as important. In this respect, the projects of Tombal[44], Assouline[58] and Tahara[54] are worthwhile following. Further large scale validation of some of these pilot projects is planned and the results should hopefully be available for RECIST 2.0.

By virtue of the search system used, it has to be assumed that this review may be incomplete. Other useful commentary may have been included in the discussion sections of phase II and phase III trial reports, which we would have missed by the methods used in searching the literature. However, we believe, we have captured many important assessments of RECIST in this review.

RECIST was implemented five years ago and since has been used at a level far beyond expectation. This indicates that it serves a purpose. The large majority of the validation studies reported in this review support the use of RECIST as a tool for tumour evaluation in most common situations. The review also identified areas requiring specific criteria or attention. Imaging in oncology, and in drug development in particular, is rapidly developing and requires continuous research and validation. A revised version of RECIST will take into account the experience accumulated so far and should provide guidance informed by the experience of many in the field to aid in assessment of novel agents.

Conflict of interest statement

The authors declare that they have no conflict of interest in relation to the work reported in this paper.

REFERENCES

1. Therasse P, Arbuck SG, Eisenhauer EA, et al. New guidelines to evaluate the response to treatment in solid tumours. *J Natl Cancer Inst* 2000;**92**:205–16.
2. Miller AB, Hogestraeten B, Staquet M, Winkler A. Reporting results of cancer treatment. *Cancer* 1981;**47**:207–14.
3. Green S, Weiss GR. Southwest oncology group standard response criteria, endpoint definitions and toxicity criteria. *Invest New Drug* 1992;**10**:239–53.
4. Tumour eligibility and response criteria for phase II and III studies. *EORTC data center manual*. Brussels; 1992. monograph.
5. Van Glabbeke M, van Oosterom AT, Steward W, Verweij J, Mouridsen H. Selection of large and objectively measurable target lesions in EORTC phase II trials: impact on recruitment and response rate. EORTC soft tissue and bone sarcoma group (STBSG). *Eur J Cancer* 1993;**29**:1943–7.
6. Padhani AR, Husband JE. Are current tumour response criteria relevant for the 21st century? *Br J Radiol* 2000;**73**:1031–3.
7. Padhani AR, Ollivier L. The RECIST (response evaluation criteria in solid tumours) criteria: implications for diagnostic radiologists. *Br J Radiol* 2001;**74**:983–6.
8. Husband JE, Schwartz LH, Spencer J, et al. Evaluation of the response to treatment of solid tumours – a consensus statement of the International Cancer Imaging Society. *Br J Cancer* 2004;**90**:2256–60.
9. Bellomi M, Preda L. Evaluation of the response to therapy of neoplastic lesions. *Radiol Med* 2004;**107**:450–8.
10. Mazumdar M, Smith A, Schwartz LH. A statistical simulation study finds discordance between WHO criteria and RECIST guideline. *J Clin Epidemiol* 2004;**57**:358–65.
11. Hillman SL, Sargent DJ, An MW, et al. Evaluation of RECIST criteria in determining the response to treatment in solid tumours: a north central cancer treatment group (NCCTG) investigation. *Proc Am Soc Clin Oncol* 2003;**22**:521. [abstr 2095].
12. Schwartz LH, Mazumdar M, Brown M, Smith A, Panicek DM. Variability in response assessment in solid tumours: effect of number of lesions chosen for measurements. *Clin Cancer Res* 2003;**9**:4318–23.
13. Mazumdar M, Smith A, Debroy PP, Schwartz LH. A theoretical approach to choosing the minimum number of multiple tumours required for assessing treatment response. *J Clin Epidemiol* 2005;**58**:150–3.
14. Perez-Garci JL, Muñoz M, Willimans G, et al. Assessment of the value of confirming responses in clinical trials in oncology. *Eur J Cancer* 2005;**41**:1528–32.
15. McHugh K, Kao S. Response evaluation criteria in solid tumours (RECIST): problems and need for modifications in paediatric oncology? *Br J Radiol* 2003;**76**:433–6.
16. Moon L, McHugh K. Advances in paediatric tumour imaging. *Arch Dis Child* 2005;**90**:608–11.
17. Barnacle AM, McHugh K. Limitations with the response evaluation criteria in solid tumours (RECIST) guidance in disseminated paediatric malignancy. *Pediatr Blood Cancer*; 2005, Oct 17 (Epub ahead of print).
18. Werner-Wasik M, Xiao Y, Pequignot E, Currant WJ, Hauck W. Assessment of lung cancer response after nonoperative therapy: tumour diameter, bidimensional product, and volume. A serial CT scan-based study. *Int J Radiat Oncol Biol Phys* 2001;**51**:56–61.
19. Cortes J, Rodriguez J, Diaz-Gonzalez JA, et al. Comparison of unidimensional and bidimensional measurements in metastatic non-small cell lung cancer. *Br J Cancer* 2002;**15**:158–60.

20. Watanabe H, Yamamoto S, Kunitoh H, et al. Tumour response to chemotherapy: the validity and reproducibility of RECIST guidelines in NSCLC patients. *Cancer Sci* 2003;**94**:1015–20.
21. Watanabe H, Leki R, Mori K, et al. The impact of an independent response evaluation committee (REC) using RECIST guidelines in a four-arm cooperative study (FACS) for advanced non-small cell lung cancer (NSCLC) in Japan. *Proc Am Soc Clin Oncol* 2004;**23**:629. (abstr 7065).
22. Erasmus JJ, Gladish GW, Broemeling L, et al. Interobserver and intraobserver variability in measurements of non-small-cell carcinoma lung lesions: implications for assessment of tumour response. *J Clin Oncol* 2003;**21**:2574–82.
23. Grossi F, Belvedere O, Fasola G, et al. Tumour measurements on computed tomographic images of non-small cell lung cancer were similar among cancer progressionals from different specialties. *J Clin Epidemiol* 2004;**57**:804–8.
24. 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.
25. Byrne MJ, Nowak AK. Modified RECIST criteria for assessment of response in malignant pleural mesothelioma. *Ann Oncol* 2004;**15**:257–60.
26. Nowak AK. CT, RECIST, and malignant pleural mesothelioma. *Lung Cancer* 2005;**49**:S37–40.
27. van Klaveren RJ, Aerts JG, de Bruin H, Giaccone G, Manegold C, van Meerbeeck JP. Inadequacy of the RECIST criteria for response evaluation in patients with malignant pleural mesothelioma. *Lung Cancer* 2004;**43**:63–9.
28. Monetti F, Casanova S, Grasso A, Cafferata MA, Ardizzoni A, Neumaie CE. Inadequacy of the new response evaluation criteria in solid tumours (RECIST) in patients with malignant pleural mesothelioma: report of four cases. *Lung Cancer* 2004;**43**:71–4.
29. Heelan R. Staging and response to therapy of malignant pleural mesothelioma. *Lung Cancer* 2004;**45**:S59–61.
30. Trillet-Lenoir V, Freyer G, Kaemmerlen P, et al. Assessment of tumour response to chemotherapy for metastatic colorectal cancer: accuracy of the RECIST criteria. *Br J Radiol* 2002;**75**:903–8.
31. Prasad SR, Saini S, Summer JE, Hahn PF, Sahani D, Boland GW. Radiological measurement of breast cancer metastases to lung and liver: comparison between WHO (bidimensional) and RECIST (unidimensional). *J Comput Assist Tomogr* 2003;**27**:380–4.
32. Choi JH, Ahn MJ, Rhim H, et al. Comparion of RECIST and WHO response criteria in metastatic colorectal carcinoma. *Proc Am Soc Clin Oncol* 2005;**23**:301. [abstr 3721].
33. Muro K, Hamaguchi T, Ohtsy A, et al. A phase II study of single-agent docetaxel in patients with metastatic esophageal cancer. *Ann Oncol* 2004;**15**:955–9.
34. Tran LN, Brown MS, Goldin JG, 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.
35. Kimura M, Tominaga T. Outstanding problems with response evaluation criteria in solid tumours (RECIST) in breast cancer. *Breast Cancer* 2002;**9**:153–9.
36. Stroobants S, Goeminne J, Seegers M, et al. 18FDG-Positron emission tomography for the early prediction of response in advanced soft tissue sarcoma treated with imatinib mesylate (Glyvec). *Eur J Cancer* 2003;**39**:2012–20.
37. Reichardt P, Schneider U, Stroszczynski C, Pink D, Hohenberger P. Molecular response of gastrointestinal stromal tumour after treatment with tyrosine kinase inhibitor imatinib mesylate. *J Clin Pathol* 2004;**57**:215–7.
38. Choi H, Charnsangavej C, de Castro Faria S, et al. CT evaluation of the response of gastrointestinal stromal tumours after imatinib mesylate treatment: a quantitative analysis correlated with FDG PET findings. *Am J Roentgenol* 2004;**183**:1619–28.
39. Goldstein D, Tan BS, Rossleigh M, Haindl W, Walker B, Dixon J. Gastrointestinal stromal tumours: correlation of F-FDG gamma camera-based coincidence positron emission tomography with CT for the assessment of treatment response – an AGITG study. *Oncology* 2005;**69**:326–32.
40. Stroszczynski C, Jost D, Reichardt P, et al. Follow-up of gastro-intestinal stromal tumours (GIST) during treatment with imatinib mesylate by abdominal MRI. *Eur Radiol* 2005;**15**:2448–56.
41. Choi H. Critical issues in response evaluation on computed tomography: lessons from the gastrointestinal stromal tumour model. *Curr Oncol Rep* 2005;**7**:307–11.
42. Therasse P, Le Cesne A, van Glabbeke M, Verweij J, Judson Ifor the EORTC Soft Tissue and Bone Sarcoma Group. RECIST vs WHO: Prospective comparison of response criteria in an EORTC phase II clinical trial investigating ET-743 in advanced soft tissue sarcoma. *Eur J Cancer* 2005;**41**:1426–30.
43. Schuetze SM, Eary JF, Griffith KA, et al. FDG PET but not RECIST agrees with histologic response of soft tissue sarcoma to neoadjuvant chemotherapy. *Proc Am Soc Clin Oncol* 2005;**23**:817s. [abstr 9005].
44. Tombal B, Rezazadeh A, Therasse P, van Cangh PH, Vande Berg B, Lecouvet FE. Magnetic resonance imaging of the axial skeleton enables objective measurement of tumour response on prostate cancer bone metastases. *Prostate* 2005;**65**:178–87.
45. Scher HI, Michael JM, Kelly WK, Schwartz LH. Prostate cancer clinical trial end points: "RECIST"ing a step backwards. *Clin Cancer Res* 2005;**11**:5223–32.
46. Warren KE, Patronas N, Aikin AA, Albert PS, Balis FM. Comparison of one-, two- and three-dimensional measurements of childhood brain tumours. *J Nat Cancer Inst* 2001;**93**:1401–5.
47. Sze G, Mehta MP, Schutlz CJ, et al. Radiologic response evaluation of brain metastases: unidimensional World Health Organization (WHO) response evaluation criteria in solid tumours (RECIST) vs bidimensional or3-dimensional criteria. *Proc Am Soc Clin Oncol* 2001;**20**:59. [abstr 234].
48. Galanis E, Maurer MJ, Ballman KV, et al. Validation of neuroradiologic response assessment in gliomas: RECIST (1D) vs 2D measurements vs computer-assisted tumour area vs volume. *Proc Am Soc Clin Oncol* 2003;**22**:106. [abstr 423].
49. Shah G, Kesari S, Xu R, et al. Comparison of 1D, 2D, 3D and volumetric parameters in measuring tumour response in high-grade gliomas in adults. *Proc Am Soc Clin Oncol* 2004;**23**:112. [abstr 1523].
50. Schwartz LH, Mazumdar M, Wang L, et al. Response assessment classification in patients with advanced renal cell carcinoma treated on clinical trials. *Cancer* 2003;**98**:1611–9.
51. Negrier S, Thiesse C, Vincent C, et al. Different measurement criteria give different tumour response classifications in patients with metastatic renal cancer (MRCC). *Proc Am Soc Clin Oncol* 2004;**23**:520. [abstr 6013].
52. Gopinath G, Ahmed A, Buscombe JR, Dickson JC, Caplin ME, Hilson AJ. Prediction of clinical outcome in treated neuroendocrine tumours of carcinoid type using functional volumes on11In-pentetreotide SPECT imaging. *Nucl Med Commun* 2004;**25**:253–7.
53. Tahara M, Ohtsu A, Hironaka S, et al. Clinical impact of criteria for complete response (CR) of primary site to treatment of esophageal cancer. *Jpn J Clin Oncol* 2005;**35**:316–23.
54. Yoshida S, Miyata Y, Ohtsu A, Boku N, Shirao K, Shimada Y. Significance of and problems in adopting response evaluation criteria in solid tumour RECIST for assessing anticancer

effects of advanced gastric cancer. *Gastric Cancer* 2000;**3**:128–33.
55. Gronlund B, Høgdall C, Hilden J, Engelholm SA, Høgdall HH, Hansen HH. Should CA-125 response criteria be preferred to response evaluation criteria in solid tumours (RECIST) for prognostication during second-line chemotherapy of ovarian carcinoma. *J Clin Oncol* 2004;**22**:4051–8.
56. Hamaoka T, Madewell JE, Podoloff DA, Hortobagyi GN, Ueno NT. Bone imaging in metastatic breast cancer. *J Clin Oncol* 2004;**22**:2942–53.
57. Sohaib SA, Turner B, Hanson JA, Farquharson M, Oliver RH, 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.
58. Assouline S, Eisenhauer E. RECIST (Response evaluation criteria in solid tumours) applied to response in lymphoma. *Proc Am Soc Clin Oncol* 2004;**23**:582. [abstr 6606].
59. Cheson BD, Horning SJ, Coiffier B, et al. Report of an international workshop to standardize response criteria for non-Hodgkin's lymphomas. NCI sponsored international working group. *J Clin Oncol* 1999;**17**:1244.
60. Park JO, Lee SI, Song SY, et al. Measuring response in solid tumours: comparison of RECIST and WHO. *Jpn J Clin Oncol* 2003;**33**:533–7.
61. Kelloff GJ, Krohn KA, Larson SM, et al. The progress and promise of molecular imaging probes in oncologic drug development. *Clin Cancer Res* 2005;**22**:7967–85.
62. Kelloff GJ, Hoffman JM, Johnson B, et al. Progress and promise of FDG-PET imaging for cancer patient management and oncologic drug development. *Clin Cancer Res* 2005;**8**:2785–808.