Leukemia (2006) 20, 1467–1473

www.nature.com/leu

LEADING ARTICLE

# International uniform response criteria for multiple myeloma

BGM Durie[1], J-L Harousseau[2], JS Miguel[3], J Bladé[4], B Barlogie[5], K Anderson[6], M Gertz[7], M Dimopoulos[8], J Westin[9], P Sonneveld[10], H Ludwig[11], G Gahrton[12], M Beksac[13], J Crowley[14], A Belch[15], M Boccadaro[16], I Turesson[17], D Joshua[18], D Vesole[19], R Kyle[7], R Alexanian[20], G Tricot[5], M Attal[21], G Merlini[22], R Powles[23], P Richardson[24], K Shimizu[25], P Tosi[26], G Morgan[27] and SV Rajkumar[7] on behalf of the International Myeloma Working Group[29]

*[1]Aptium Oncology, Inc., Cedars-Sinai Outpatient Cancer Center, Los Angeles, CA, USA; [2]Institute de Biologie, Nantes, France; [3]University of Salamanca, Salamanca, Spain; [4]Hospital Clinica, Barcelona, Spain; [5]MIRT UAMS, Little Rock, Arkansas, USA; [6]DFCI, Boston, MA, USA; [7]Mayo Clinic, Rochester, MN, USA; [8]Alexandra Hospital, Athens, Greece; [9]University of Gothenberg, Gothenberg, Sweden; [10]Rotterdam, The Netherlands; [11]Wilhelminenspital Der Stat Wien, Vienna, Austria; [12]Karolinska Institutet, Stockholm, Sweden; [13]Ankara University, Turkey; [14]Cancer Research and Biostatistics, Seattle, WA, USA; [15]Cross Cancer Institute, Canada; [16]University of Torino, Torino, Italy; [17]University of Malmo, Malmo, Sweden; [18]Royal Prince Alfred Hospital, Sydney, Australia; [19]St Vincent's Comprehensive Cancer Center, New York, NY, USA; [20]MD Anderson, Houston, TX, USA; [21]Purpan Hospital, Toulouse, France; [22]University of Pavia, Pavia, Italy; [23]The Leukemia and Myeloma Program, Wimbledon, UK; [24]Dana Farber Cancer Institute, Boston, MA, USA; [25]Nagoya City Midori General Hospital, Nagoya, Japan; [26]University of Bologna, Bologna, Italy and [27]Royal Marsden Hospital, London, UK*

**New uniform response criteria are required to adequately assess clinical outcomes in myeloma. The European Group for Blood and Bone Marrow Transplant/International Bone Marrow Transplant Registry criteria have been expanded, clarified and updated to provide a new comprehensive evaluation system. Categories for stringent complete response and very good partial response are added. The serum free light-chain assay is included to allow evaluation of patients with oligo-secretory disease. Inconsistencies in prior criteria are clarified making confirmation of response and disease progression easier to perform. Emphasis is placed upon time to event and duration of response as critical end points. The requirements necessary to use overall survival duration as the ultimate end point are discussed. It is anticipated that the International Response Criteria for multiple myeloma will be widely used in future clinical trials of myeloma.**
*Leukemia* (2006) **20,** 1467–1473. doi:10.1038/sj.leu.2404284; published online 20 July 2006
**Keywords:** myeloma; response criteria; staging; survival duration; uniform criteria; clinical outcomes

## Introduction

There is an increasing need for widely accepted, reproducible criteria to evaluate response in multiple myeloma.[1,2] Several different systems are currently in use, but are not exactly comparable. For example, the US cooperative groups ECOG and SWOG have differing systems, as do several European groups, such as the MRC (UK)[3] and the IFM (France).[4] In addition, the European Group for Blood and Bone Marrow Transplant/International Bone Marrow Transplant Registry/American Bone Marrow Transplant Registry (EBMT/IBMTR/ABMTR) developed widely used criteria, commonly referred to as the EBMT criteria.[5] However, as discussed below, there is a need to update prior criteria.

Correspondence: Dr BGM Durie, Aptium Oncology Inc., Cedars-Sinai Outpatient Cancer Center, 8201 Beverly Boulevard, Los Angeles, CA 90048, USA or Dr SV Rajkumar, Division of Hematology, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, USA.
E-mails: bdurie@aptiumoncology.com or rajkumar.vincent@mayo.edu.
[29]For members of the International Myeloma Working Group, see Appendix A.
Received 8 May 2006; accepted 11 May 2006; published online 20 July 2006

The need for new uniform response criteria has been triggered by several factors (Table 1). The most pressing need is for criteria that facilitate precise comparisons between new treatment strategies. Better criteria are also required for use in the clinic at the individual patient level. In this setting, clarification of complete response (CR) is particularly important. As more active agents are available, there is a need to assess not just if response has occurred, but the exact magnitude of response. There is increased awareness of the distinction between surrogate end points such as reduction in M-component level and more clinical end points such as recovery of functional status or organ function, length of response and overall survival duration.[6]

Many of the commonly used criteria do not define CR stringently. In the EBMT criteria, CR does not require absence of monoclonal (M) plasma cells, but rather the reduction in plasma cells to 5% or less on bone marrow samples. This naturally results in the contamination of a subset of complete responders with normal polyclonal plasma cells in the marrow with those who still have M plasma cells. The latter are easily detected by kappa/lambda immunostaining or by immunofluorescence studies using flow cytometry. Specific categories of CR with varying degrees of stringency allow greater precision in the definition of CR, enable comparison of the efficacy of various treatments including novel agents and can permit the detection and monitoring of relapse more accurately. Existing criteria lack sufficient detail, which as a result allows substantial investigator discretion, and leads to inaccuracies in the estimated response rate. For example, the EBMT criteria require specific reductions in M-protein levels for each category of response, but the minimum level of M-protein that is required in the serum and urine to allow accurate response assessment is not specified. Similarly, it is not clear from prior criteria how patients with 'unmeasurable' levels of urine M-protein should be monitored for response evaluation.

Finally, present criteria allow limited assessment of response in patients with oligo-secretory or non-secretory myeloma.[7] Response in these patients can now be assessed using the sensitive serum free light-chain (FLC) assay (Freelite, Binding Site). Incorporation of the serum FLC assay into the response criteria for myeloma allows inclusion and evaluation of these patients in clinical trials.

**Table 1** Rationale for the development of uniform response criteria

- Facilitate precise comparisons of efficacy between new treatment strategies in trials
- Incorporation of the serum FLC assay to include assessment of patients with oligo-secretory and non-secretory disease[a]
- Stricter definitions for CR
- Provide clarifications, improve detail and correct inconsistencies in prior response criteria

Abbreviations: CR, complete response; FLC, free light chain.
[a]Oligo- and non-secretory myeloma identifies patients without sufficient M-component in serum and/or urine to monitor response (see Table 4: definitions of measurable disease).

## Development of new response and relapse criteria

The International Myeloma Working Group has developed new standard diagnostic criteria[8] and the new International Staging System (ISS) for multiple myeloma,[9] which are being widely accepted as the current standards for diagnosis and staging. The development of the new response criteria proposed in this manuscript started with a meeting of the International Myeloma Working Group (participants are listed at the end of the manuscript) during the 10th International Myeloma Workshop, Sydney, 10–14 April 2005. Based on the discussions and decisions made at this meeting, the criteria were formulated and drafted by two of the authors (BGD and SVR) and circulated to the members of the Working Group and revised. Final approval was made at a meeting of the International Myeloma Working Group at the Annual Meeting of the American Society of Hematology, Atlanta, GA, USA, December 2005 and subsequent reviews of this paper.

A summary of the important changes in the new criteria versus prior systems is provided in Table 2. It is important to point out that for patients with measurable M-protein levels in the serum and urine, the definitions of complete and partial response as well as disease progression match those used in the EBMT (Bladé) criteria. Therefore, although important clarifications are added, for all practical purposes, in trials that include only patients with measurable disease response rates and progression, estimates reported using the using the new International Myeloma Working Group criteria will be comparable to those using the EBMT criteria. This will allow easy comparison of rates reported in trials using the EBMT criteria with those using the new criteria. The most important changes in the new criteria are (1) addition of a new category of stringent CR that is of significant importance given rapid advances in therapy, (2) addition of response criteria for interpreting the serum FLC assay, which will enable numerous patients hitherto excluded from clinical trials for lack of measurable disease to enter and be evaluated on clinical trials, and (3) formal addition of a category of very good partial response (VGPR) to allow distinction of patients with excellent responses that may have outcomes similar to those patients considered to be in CR.

## Diagnostic criteria for multiple myeloma

The need for clear baseline diagnostic criteria cannot be overemphasized. Three recent publications from the International Myeloma Working Group incorporate recommended methods for diagnosis, baseline staging and prognostic classifications as well as disease subtype identification.[1,8,9] With these systems, the features of patients entering clinical trials can be clearly delineated. Table 3 summarizes the diagnostic criteria for active myeloma.

**Table 2** Summary of similarities and specific changes introduced in the New Uniform Response Criteria compared to the EBMT/IBMTR Criteria

- For patients with measurable levels of serum and urine monoclonal protein levels, the criteria for CR, PR and progressive disease remain unchanged. (Tables 5 and 6)
- Clarification and revision of important practical details of response evaluation (Table 4)
  - Elimination of mandatory 6 weeks wait time to confirm achievement of response
  - Introduction of a similar non-time-dependent confirmation for relapse and/or disease progression
  - Clarification of the 'start time' for duration of response evaluation
  - Requirement of $\geqslant$PR as response requirement for new drug trials
  - Allow use of quantitative immunoglobulin levels in patients in whom the M-protein measurements are unavailable or unreliable
- Introduction of new response categories (Table 5) sCR and VGPR
  - Elimination of the minor response category
- Incorporation of response criteria for the serum FLC assay to enable assessment of response in patients with non- or oligo-secretory disease (Tables 5 and 6)
- Clarification that criteria for progressive disease (rather than criteria for 'relapse from CR') are to be used for calculation of time to progression and progression-free survival in patients who are in CR. Criteria for relapse from CR are to be used only if DFS is calculated and reported
- Introduction of new category of clinical relapse or progressive disease (Table 6)
  - Introduces clinical relapse as a new optional end point

Abbreviations: CR, complete response; DFS, disease-free survival; EBMT, European Group for Blood and Bone Marrow Transplant; FLC, free light chain; IBMTR, International Bone Marrow Transplant Registry; PR, partial response; sCR, stringent complete response; VGPR, very good partial response.

## Response categories

Changes in the M-component level are the principal indicators used for response evaluation.[6] It is important to note that M-component is a surrogate marker and its use is accompanied by all the pitfalls that can potentially detract from such use including variations in marker synthesis, metabolism or release as well as myeloma cell heterogeneity with respect to M-component production.[10] The major response categories include CR, partial response (PR), stable disease (SD), progressive disease (PD) and relapse from CR (see Tables 5 and 6).

Additional subcategories have been used by a number of investigators.[3,11] The subcategories of near complete response (nCR) and VGPR have been integrated into the new criteria under one single category termed 'VGPR'. Importantly, the term 'stable disease' is not recommended for use as a measure of treatment efficacy; instead time to progression (TTP) and response duration estimates (see below) should be used in instances when the stability of disease with a particular therapy needs to be highlighted. TTP is calculated from the start of treatment and includes all patients entering the trial. Duration of response (DOR) is calculated from the time of first recorded achievement of a particular response level, that is, PR, VGPR, CR or sCR (see Table 5), and includes only responding patients. Although documentation of response requires a confirmatory measurement, the start time for DOR is the first date at which response was noted.

**Table 3** Diagnostic criteria for multiple myeloma requiring systemic therapy

Presence of an M-component[a] in serum and/or urine plus clonal plasma cells in the bone marrow and/or a documented clonal plasmacytoma

PLUS one or more of the following:[b]
- Calcium elevation (>11.5 mg/dl) [>2.65 mmol/l]
- Renal insufficiency (creatinine >2 mg/dl) [177 $\mu$mol/l or more]
- Anemia (hemoglobin <10 g/dl or 2 g/dl <normal) (hemoglobin <12.5 mmol/l[c] or 1.25 mmol/l<normal)
- Bone disease (lytic lesions or osteopenia)

[a]In patients with no detectable M-component, an abnormal serum FLC ratio on the serum FLC assay can substitute and satisfy this criterion. For patients, with no serum or urine M-component and normal serum FLC ratio, the baseline bone marrow must have ⩾10% clonal plasma cells; these patients are referred to as having 'non-secretory myeloma'. Patients with biopsy-proven amyloidosis and/or systemic light chain deposition disease (LCDD) should be classified as 'myeloma with documented amyloidosis' or 'myeloma with documented LCDD,' respectively if they have ⩾30% plasma cells and/or myeloma-related bone disease.
[b]Must be attributable to the underlying plasma cell disorder.
[c]*Note*: Hemoglobin of 10 g/dl is 12.5 mmol/l [or 100 g/l].

## Important aspects of response assessment

Table 4 summarizes important practical details in response assessment. Two specific points must be emphasized. Firstly, checking the M-component level at each cycle during induction is critically important in the evaluation of novel therapies to determine the speed of response, which may have clinical implications. For example, with several new regimens, response occurs rapidly and can be substantial within 1–2 months.[12,13] The second point is that the new criteria eliminate the need for consecutive confirmations 6 weeks apart currently required for response testing. A DOR of 6 weeks does not carry major clinical significance and is not a surrogate for durability of response. The main concern is to eliminate laboratory or other error; this can be carried out by the requirement of a confirmatory test at any time following the first test provided it is before any new/non-protocol therapy. The importance of response, that is, its durability, should be highlighted by reporting data on TTP and DOR. Thus plateau phase can be documented by indicating the TTP and/or DOR.[1]

Three aspects pertaining to the serum FLC assay deserve emphasis. First, the serum FLC assay (Freelite, The Binding Site, Birmingham, UK) is a highly sensitive marker of light chains in circulation that are unbound to intact immunoglobulin, and the FLC ratio is an excellent indicator of clonality.[14] Thus, normalizing of serum FLC ratio is a stricter indicator of CR, and may correlate well with extended response duration[15] (Kumar S *et al*. *Blood* 2005; **106**: 971a, abs 3479). Note that in patients with renal insufficiency, the levels of both the kappa and lambda may remain elevated, but the ratio normalizes with achievement of CR. Second, in order to minimize chance of error, FLC response is not assessable for patients who start with low baseline serum FLC assay levels below 10 mg/dl (<100 mg/l). Third, although the serum FLC assay is a very reliable test, it is important to closely monitor laboratory variation.[16] Strict guidelines are required with regard to usage times for the serum FLC assay kits. It should also be noted that serum FLC assay testing might be useful in the prognostic and response evaluation of patients who also have a measurable serum and/or urine M-component in the future, given its recently reported prognostic value in M-gammopathy of undetermined significance (MGUS).[17]

**Table 4** Practical details of response evaluation

*Laboratory tests for measurement of M-protein*
- Serum M-protein level is quantitated using densitometry on SPEP except in cases where the SPEP is felt to be unreliable such as in patients with IgA monoclonal proteins migrating in the beta region. If SPEP is not available or felt to be unreliable (e.g., in some cases of IgA myeloma) for routine M-protein quantitation during therapy, then quantitative immunoglobulin levels on nephelometry or turbidometry can be accepted. However, this must be explicitly reported, and only nephelometry can be used for that patient to assess response and SPEP and nephelometric values cannot be used interchangeably.
- Urine M-protein measurement is estimated using 24-h UPEP only. Random or 24 h urine tests measuring kappa and lambda light chain levels are not reliable and are not recommended

*Definitions of measurable disease*
- Response criteria for all categories and subcategories of response except CR are applicable only to patients who have 'measurable' disease defined by at least one of the following three measurements:
  Serum M-protein ⩾1 g/dl (⩾10 gm/l)[10 g/l]
  Urine M-protein ⩾200 mg/24 h
  Serum FLC assay: Involved FLC level ⩾10 mg/dl (⩾100 mg/l) provided serum FLC ratio is abnormal

  *Response criteria for CR* are applicable for patients who have abnormalities on one of the three measurements. Note that patients who do not meet any of the criteria for measurable disease as listed above can only be assessed for stringent CR, and cannot be assessed for any of the other response categories

*Follow-up to meet criteria for PR or SD*
- It is recommended that patients undergoing therapy be tracked monthly for the first year of new therapy and every other month thereafter
- Patients with 'measurable disease' as defined above need to be followed by both SPEP and UPEP for response assessment and categorization
- Except for assessment of CR, patients with measurable disease restricted to the SPEP will need to be followed only by SPEP; correspondingly, patients with measurable disease restricted to the UPEP will need to be followed only by UPEP[a]
- Patients with measurable disease in either SPEP or UPEP or both will be assessed for response only based on these two tests and not by the FLC assay. FLC response criteria are only applicable to patients without measurable disease in the serum or urine, and to fulfill the requirements of the category of stringent CR
- To be considered CR, both serum and urine immunofixation must be carried out and be negative regardless of the size of baseline M-protein in the serum or urine; patients with negative UPEP values pretreatment still require UPEP testing to confirm CR and exclude light chain or Bence–Jones escape
- Skeletal survey is not required for assessment of response unless clinically indicated, but is recommended once a year in clinical practice; bone marrow is required only for categorization of CR, and for patients with non-secretory disease

Abbreviations: CR, complete response; FLC, free light chain; PR, partial response; SD, stable disease; SPEP, serum protein electrophoresis; UPEP, urine protein electrophoresis.
[a]For good clinical practice patients should be periodically screened for light chain escape with UPEP or serum FLC assay.

## The international Myeloma Working Group Uniform Response Criteria

The International Myeloma Working Group Uniform Response Criteria are listed in Table 5. Under CR two categories are listed: CR and stringent (sCR). The CR category is available for widespread use and provides continuity with prior systems. However, sCR, the more stringent category, allows more

**Table 5** International Myeloma Working Group uniform response criteria: CR and other response categories

| *Response subcategory* | *Response criteria*[a] |
|---|---|
| sCR | CR as defined below plus<br>Normal FLC ratio and<br>Absence of clonal cells in bone marrow[b] by immunohistochemistry or immunofluorescence[c] |
| CR | Negative immunofixation on the serum and urine and<br>Disappearance of any soft tissue plasmacytomas and<br>$\leqslant$5% plasma cells in bone marrow[b] |
| VGPR | Serum and urine M-protein detectable by immunofixation but not on electrophoresis or 90% or greater reduction in serum M-protein plus urine M-protein level $<$100 mg per 24 h |
| PR | $\geqslant$50% reduction of serum M-protein and reduction in 24-h urinary M-protein by $\geqslant$90% or to $<$200 mg per 24 h<br>If the serum and urine M-protein are unmeasurable,[d] a $\geqslant$50% decrease in the difference between involved and uninvolved FLC levels is required in place of the M-protein criteria<br>If serum and urine M-protein are unmeasurable, and serum free light assay is also unmeasurable, $\geqslant$50% reduction in plasma cells is required in place of M-protein, provided baseline bone marrow plasma cell percentage was $\geqslant$30%<br>In addition to the above listed criteria, if present at baseline, a $\geqslant$50% reduction in the size of soft tissue plasmacytomas is also required |
| SD (not recommended for use as an indicator of response; stability of disease is best described by providing the time to progression estimates) | Not meeting criteria for CR, VGPR, PR or progressive disease |

Abbreviations: CR, complete response; FLC, free light chain; PR, partial response; SD, stable disease; sCR, stringent complete response; VGPR, very good partial response.
[a]All response categories require two consecutive assessments made at anytime before the institution of any new therapy; all categories also require no known evidence of progressive or new bone lesions if radiographic studies were performed. Radiographic studies are not required to satisfy these response requirements.
[b]Confirmation with repeat bone marrow biopsy not needed.
[c]Presence/absence of clonal cells is based upon the $k/\lambda$ ratio. An abnormal $k/\lambda$ ratio by immunohistochemistry and/or immunofluorescence requires a minimum of 100 plasma cells for analysis. An abnormal ratio reflecting presence of an abnormal clone is $k/\lambda$ of $>$4:1 or $<$1:2.
[d]Refer to Table 4 for definitions of measurable disease.

accurate assessment of new therapies. Many myeloma groups already use this latter category. It is now possible to specifically list and clearly identify which categories are used. The major goal is to foster studies evaluating correlations between stringent CR and durable response and prolonged survival.

VGPR, as defined in IFM trials,[4] has been very slightly modified to also include what has been called nCR. Use of VGPR has several advantages including the reliance upon the 90% or higher regression cutoff, which is simpler to implement than use of immunofixation positivity versus negativity, an observer-dependent assessment. In addition, failure to achieve VGPR correlates with inferior outcome.[1] The definition of PR except for inclusion of the FLC assay for the subgroup of patients with 'unmeasurable' disease is similar to the EBMT criteria. It is important to note that the FLC assay should not be used to assess response in patients with measurable levels of M-protein in either serum or urine. Such patients should be assessed using standard criteria; the serum light-chain assay is only applicable to those patients who do not have either 1 g/dl or higher M-protein in the serum or 200 mg/day or higher M-protein level in the urine. Less than PR is identified as SD, which can be clinically meaningful, but is not sufficient as an indicator of response benefit in new therapeutic trials. Reporting SD or response categories less than PR as meaningful is not recommended in clinical trials of new agents. Overall, the emphasis is upon simplicity, reproducibility and the awareness that very fine discriminations are frequently unreliable and not clinically meaningful.

The criteria for PD and relapse from CR are listed in Table 6. A category of clinical relapse has been added for optional assessment in clinical trials and for use in clinical practice. Progressive disease will continue to identify patients in whom the standard M-component (and related) criteria for relapse or disease progression have been met. Progressive disease is the end point that is used for calculating TTP and progression-free survival (PFS) in trials, and mirrors the EBMT criteria. One problem is that progression defined using these criteria may or may not reflect a need for therapy (or new therapy). Early re-treatment can be unnecessary, result in unwanted toxicities and underestimate the benefit of prior treatment, as true symptomatic relapse may not emerge until months or years later. Thus, discrete 'event categories' are required to identify relapse or progression requiring intervention. These 'events' are broadly the same as the CRAB categories used for diagnosis of myeloma. Various nuances and details related to use in the relapse setting are outlined in Table 6. Thus, where possible, reporting of time to re-treatment and/or time to clinical relapse would be useful; as mentioned earlier, these definitions will also be useful in clinical practice.

The difficulties and nuances in evaluating myeloma-related events are well known. It is important to re-emphasize that myeloma must be the cause of events. Whatever additional testing is required to confirm myeloma relatedness is strongly encouraged. This may include magnetic resonance imaging, computed tomography and/or fluoro-18-deoxyglucose (FDG)/positron emission tomography (PET) imaging (Walker R *et al*. *Blood* 2004; **104**: 217a, abs 758).[18,19]

**Table 6** International Myeloma Working Group uniform response criteria: disease progression and relapse

| *Relapse subcategory* | *Relapse criteria* |
|---|---|
| Progressive disease[a]<br>To be used for calculation of time to progression and progression-free survival end points for all patients including those in CR (includes primary progressive disease and disease progression on or off therapy) | Progressive Disease: requires any one or more of the following:<br><br>Increase of ⩾25% from baseline in<br>Serum M-component and/or (the absolute increase must be ⩾0.5 g/dl)[b]<br>Urine M-component and/or (the absolute increase must be ⩾200 mg/24 h<br>Only in patients without measurable serum and urine M-protein levels: the difference between involved and uninvolved FLC levels. The absolute increase must be >10 mg/dl.<br>Bone marrow plasma cell percentage: the absolute % must be ⩾10%[c]<br>Definite development of new bone lesions or soft tissue plasmacytomas or definite increase in the size of existing bone lesions or soft tissue plasmacytomas<br>Development of hypercalcemia (corrected serum calcium >11.5 mg/dl or 2.65 mmol/l) that can be attributed solely to the plasma cell proliferative disorder |
| Clinical relapse[a] | Clinical relapse requires one or more of:<br>Direct indicators of increasing disease and/or end organ dysfunction (CRAB features)[b] It is not used in calculation of time to progression or progression-free survival but is listed here as as something that can be reported optionally or for use in clinical practice<br>1. Development of new soft tissue plasmacytomas or bone lesions<br>2. Definite increase in the size of existing plasmacytomas or bone lesions. A definite increase is defined as a 50% (and at least 1 cm) increase as measured serially by the sum of the products of the cross-diameters of the measurable lesion<br>3. Hypercalcemia (>11.5 mg/dl) [2.65 mmol/l]<br>4. Decrease in hemoglobin of ⩾2 g/dl [1.25 mmol/l] (see Table 3 for further details)<br>5. Rise in serum creatinine by 2 mg/dl or more [177 $\mu$mol/l or more] |
| Relapse from CR[a](To be used only if the end point studied is DFS)[d] | Any one or more of the following:<br>Reappearance of serum or urine M-protein by immunofixation or electrophoresis<br>Development of ⩾5% plasma cells in the bone marrow[c]<br>Appearance of any other sign of progression (i.e., new plasmacytoma, lytic bone lesion, or hypercalcemia see below) |

Abbreviations: CR, complete response; DFS, disease-free survival.
[a]All relapse categories require two consecutive assessments made at anytime before classification as relapse or disease progression and/or the institution of any new therapy.
[b]For progressive disease, serum M-component increases of ⩾1 gm/dl are sufficient to define relapse if starting M-component is ⩾5 g/dl.
[c]Relapse from CR has the 5% cutoff versus 10% for other categories of relapse.
[d]For purposes of calculating time to progression and progression-free survival, CR patients should also be evaluated using criteria listed above for progressive disease.

## Survival end points

End points such as PFS, TTP and DOR can predict ultimate overall survival (Tricot G *et al*. *Blood* 2004; **104**: 265a, abs 926).[6,20–22] Several different methods are used to calculate response duration and the impact of treatment.

- *PFS*: PFS is the time from start of the treatment to disease progression or death. This encompasses all patients and has been considered a surrogate marker for overall survival duration. This is the recommended method to present trial results.
- *Event-free survival (EFS)*: The definition for EFS depends on how 'event' is defined. In some studies, this can be the same as PFS. EFS can also include additional 'events' that are considered to be of importance besides death, including serious drug toxicity. EFS is not recommended for general use unless specifically defined, as confusion can arise about the details of additional 'events'. PFS is preferred.
- *TTP*: This is the time from start of treatment to disease progression with deaths owing to causes other than progression not counted, but censored. This is a helpful method to assess the durability of treatment benefit.
- *Disease-free survival (DFS)*: DFS applies to patients in CR, and is measured from the start of CR to the time of relapse from CR. This parameter has limited value in myeloma at present.
- *DOR*: DOR applies to patients achieving at least PR by the criteria in Table 5, and is measured from start of achieving PR (first observation of PR before confirmation) to the time of disease progression, with deaths owing to causes other than progression not counted, but censored. This is an additional parameter for consideration in the assessment of new agents and/or new comprehensive treatment strategies. DOR and TTP are the recommended ways of establishing the durability of response.

## Overall survival

Many recent myeloma trials have had response and/or TTP as the primary end points. However, overall survival and quality of life reflect the full impact of therapies. Several factors limit the use of overall survival as the ultimate end point.

- Over 5 years of follow-up are required to assess benefit.
- Initial response and TTP may or may not translate into overall survival benefit.
- New agents used as part of induction, consolidation/transplant and/or maintenance are frequently used at time of relapse in the 'control' (non-use) arm of trials. Thus the comparison is with early versus later use. There has been no widely accepted plan or framework to control for this.

- Additional new agents are now being introduced, which can further impact outcome assessment.

The problems involved are illustrated by several trials.[23–29] In a recent trial reported by the Arkansas group,[27] thalidomide was used as part of the TT-2 in one arm of the trial and produced a significantly higher CR rate and disease-free interval. However, overall survival was *not* improved. But, it is important to note that 83% of patients not in the thalidomide arm received thalidomide at relapse. Thus, the study reflects an unplanned 'early' versus 'later' use of a therapeutic intervention, in this case thalidomide. In a more minor way, this was also an issue in the recently published[28] results of the melphalan/prednisone (MP) versus MP thalidomide trial. New trial designs to evaluate survival duration must accommodate these types of complexity. These details are further discussed in a recent review.[1]

## Conclusions

The response criteria outlined in this paper are expected to be used widely in future clinical trials of myeloma. The major new additions to the response criteria are categories of stringent CR, VGPR and incorporation of the serum FLC assay to evaluate patients with oligo-secretory disease. The criteria also clarify several inconsistencies in prior response criteria, make confirmation of response and disease progression easier to perform with less chance of deviations, and define time to event end points that are critical in the evaluation of outcome.

## References

1 Durie BGM. New approaches to treatment for multiple myeloma: durable remission and quality of life as primary goals. *Clin Lymphoma Myeloma* 2005; **6**: 181–190.
2 Rajkumar SV, Dispenzieri A. Evaluation and monitoring of response to therapy in multiple myeloma. *Haematologica* 2005; **90**: 1305–1308.
3 Child JA, Morgan G, Davies F, Owen RG, Bell SE, Hawkins K *et al.* High-dose chemotherapy with hematopoietic stem cell rescue for multiple myeloma. *N Engl J Med* 2003; **348**: 1875–1883.
4 Attal M, Harousseau JL, Facon T, Guilhot F, Doyen C, Fuzibet JG *et al.* Single versus double autologous stem-cell transplantation for multiple myeloma. *N Engl J Med* 2003; **26**: 2495–2502.
5 Blade J, Samson D, Reece D, Apperley J, Bjorkstrand B, Gahrton G *et al.* Criteria for evaluating disease response and progression in patients with multiple myeloma treated by high-dose therapy and haemopoietic stem cell transplantation. *Br J Haematol* 1998; **102**: 1115–1123.
6 Durie BGM, Jacobson J, Barlogie B, Crowley J. Magnitude of response with myeloma frontline therapy does not predict outcome: importance of time to progression in Southwest Oncology Group chemotherapy trials. *J Clin Oncol* 2004; **22**: 1857–1863.
7 Dispenzieri A, Kyle RA, Lacy MQ, Rajkumar SV, Therneau TM, Larson DR *et al.* POEMS syndrome: definitions and long-term outcome. *Blood* 2003; **101**: 2496–2506.
8 International Myeloma Working Group. Criteria for the classification of monoclonal gammopathies, multiple myeloma and related disorders: a report of the International Myeloma Working Group. *Br J Haematol* 2003; **121**: 749–757.
9 Greipp PR, San Miguel J, Durie BGM, Crowley J, Barlogie B, Blade J *et al.* International staging system for multiple myeloma. *J Clin Oncol* 2005; **23**: 3412–3420.
10 Durie BGM. Staging and kinetics in multiple myeloma. In: Salmon SE (ed). *Clinics in Haematology: Myeloma and Related Disorders,* Vol. 2, Number 1, WB Saunders Company, Ltd: Alden Press, Ocford, 1982, pp 3–18.
11 Richardson PG, Barlogie B, Berenson J, Singhal S, Jagannath S, Irwin D *et al.* A phase II study of bortezomib in relapsed, refractory myeloma. *N Engl J Med* 2003; **348**: 2609–2617.
12 Oakervee HE, Popat R, Curry N, Smith P, Morris C, Drake M *et al.* PAD combination therapy (PS-341/bortezomib, doxorubicin and dexamethasone) for previously untreated patients with multiple myeloma. *Br J Haematol* 2005; **129**: 755–762.
13 Drayson M, Tang LX, Drew R, Mead GP, Carr-Smith H, Bradwell AR. Serum free light chain measurements for identifying and monitoring patients with nonsecretory multiple myeloma. *Blood* 2001; **97**: 2900–2902.
14 Bradwell AR, Carr-Smith HD, Mead GP, Harvey TC, Drayson MT. Serum testing for assessment of patients with Bence–Jones myeloma. *Lancet* 2003; **361**: 489–491.
15 Mead GP, Carr-Smith HD, Drayson MT, Morgan GJ, Child A, Bradwell AR. Serum free light chains for monitoring multiple myeloma. *Br J Haematol* 2004; **126**: 348–354.
16 Hassoun H, Reich L, Klimek VM, Dhodapkar M, Cohen A, Kewalramani T *et al.* Doxorubicin and dexmaethasone followed by thalidomide and dexamethasone is an effective well tolerated initial therapy for multiple myeloma. *Br J Hematol* 2006; **132**: 155.
17 Rajkumar SV, Kyle RA, Therneau TM, Melton III LJ, Bradwell AR, Clark RJ *et al.* Serum free light chain ratio is an independent risk factor for progression in monoclonal gammopathy of undetermined significance. *Blood* 2005; **106**: 812–817.
18 Kyle RA, Schreiman JS, McLeod RA, Beabout JW. Computed tomography in diagnosis and management of multiple myeloma and its variants. *Arch Int Med* 1985; **145**: 1451–1452.
19 Durie BGM, Waxman AD, D'Angeloa A, William CM. Whole body F-FDG$^{18}$ PET identifies high-risk myeloma. *J Nucl Med* 2002; **43**: 1457–1463.
20 Durie BGM. Is magnitude of initial response predictive for survival in multiple myeloma? *Ann Oncol* 1999; **2**: 166–169.
21 Lahuerta JJ, Martinez-Lopez J, de la Serna J, Blade J, Grande C, Alegra A *et al.* Remission status defined by immunofixation versus electrophoresis following autologus transplantation has a major impact on the outcome of multiple myeloma patients. *Br J Haematol* 2000; **109**: 438–446.
22 Riccardi A, Mora O, Tinelli C, Porta C, Danova M, Brugnatelli S *et al.* Response to first-line chemotherapy and long-term survival in patients with multiple myeloma: results of the MM87 prospective randomized protocol. *Eur J Cancer* 2003; **39**: 31–37.
23 Cavo M, Zamagni E, Tosi P, Tacchetti P, Cellini C, Cangini D *et al.* Superiority of thalidomide and dexamethasone over vincristine–doxorubicin–dexmaethasone (VAD) as primary therapy in preparation for autologous transplantation for multiple myeloma. *Blood* 2005; **106**: 359.
24 Richardson PG, Sonneveld P, Schuster MW, Irwin D, Stadtmauer E, Facon T *et al.* Bortezomib or high-dose dexamethasone for relapsed multiple myeloma. *N Engl J Med* 2005; **352**: 2487–2498.
25 Rajkumar SV, Dispenzieri A. Evaluation and monitoring of response to therapy in multiple myeloma. *Haematologica* 2005; **90**: 1305–1308.
26 Barlogie B, Kyle RA, Anderson KC, Greipp PR, Lazarus HM, Hurd DD *et al.* Standard chemotherapy compared with high dose chemoradiotherapy for multiple myeloma: final results of phase III US Intergroup trial S9321. *J Clin Oncol* 2006; **24**: 929–936.
27 Barlogie B, Tricot G, Anaissie E, Shaughnessy J, Rasmussen E, van Rhee F *et al.* Thalidomide and hematopoietic-cell transplantation for multiple myeloma. *N Engl J Med* 2006; **354**: 1021–1030.
28 Palumbo A, Bringhen S, Caravita T, Merla E, Capparella V, Callea V *et al.* Oral melphalan and prednisone therapy plus thalidomide compared with melphalan and prednisone alone in elderly patients with multiple myeloma: a randomized controlled trial. *Lancet* 2006; **367**: 825–831.
29 Alexanian R, Weber D, Giralt S, Dimopoulos M, Delasalle K, Smith T *et al.* Impact of complete remission with intensive therapy in patients with responsive multiple myeloma. *Bone Marrow Transplant* 2001; **27**: 1037–1043.

**Appendix A.**

*International myeloma working group:*

Raymond Alexanian, MD Anderson Center, Houston, Texas, USA
Kenneth Anderson, DFCI, Boston, Massachusetts, USA
Michael Attal, Purpan Hospital, Toulouse, France
Herve Avet-Loiseau, Institute de Biologie, Nantes, France
Leif Bergsagel, Mayo Clinic Scottsdale, Arizona, USA
Joan Bladé, Hospital Clinica, Barcelona, Spain
Bart Barlogie, MIRT UAMS Little Rock, Arkansas, USA
Regis Batille, Institute de Biologie, Nantes, France
Meral Beksac, Ankara University, Turkey
Andrew Belch, Cross Cancer Institute, Canada
Mario Boccadoro, University of Torino, Italy
Michele Cavo, Bologna, Italy
Tony Child, Leeds General Hospital, Leeds, United Kingdom
Ray Comenzo, Memorial-Sloan Kettering Cancer Center, New York, USA
John Crowley, Cancer Research and Biostatistics, Seattle, Washington, USA
William Dalton, H Lee Moffitt, Tampa, Florida, USA
Faith Davies, Royal Marsden Hospital, London, England
Meletios Dimopoulos, Alexandra Hospital, Athens, Greece
Angela Dispenzieri, Mayo Clinic, Rochester, Minnesota, USA
Brian Durie, Cedars-Sinai Outpatient Medical Center, Los Angeles, California, USA
Thierry Facon, Centre Hospitalier Regional Universitaire de Lille, France
Dorotea Fantl, Hospital Italiano, Buenos Aires, Argentina
Jean-Paul Fermand, Paris, France
Rafael Fonseca, Mayo Clinic Scottsdale, Scottsdale, Arizona, USA
Gosta Gahrton, Karolinska Institutet, Stockholm , Sweden
Morie Gertz, Mayo Clinic, Rochester, Minnesota, USA
Hartmut Goldschmidt, University of Heidelberg, Heidelberg, Germany
Philip Greipp, Mayo Clinic, Rochester, Minnesota, USA
Roman Hajek, Brno University, Brno, Czech Republic
Jean-Luc Harousseau, Institute de Biologie, Nantes, France
Kim Hawkins, University of Leeds, United Kingdom
Martin Hjorth, Rotterdam, The Netherlands
Vania Hungria, Clinica San Germano, Sao Paolo, Brazil
Mohamad Hussein, The Cleveland Clinic, Cleveland, Ohio, USA
Peter Jacobs, South Africa
Sundar Jagannath, St Vincent's Comprehensive Cancer Center, New York, USA
Mariana Juni, Fundaleu, Buenos Aires, Argentina
Douglas Joshua, Royal Prince Alfred Hospital, Sydney, Australia
Michael Katz, International Myeloma Foundation, Los Angeles, California, USA
Michio Kawano, Yamaguchi University, Ube, Japan
Shaji Kumar, Mayo Clinic, Rochester, Minnesota, USA
Robert Kyle, Mayo Clinic, Rochester, Minnesota, USA
Juan Lahuerta, Madrid, Spain
Henk Lokhorst, University Hospital, Utrecht, The Netherlands
Heinz Ludwig, Wilhelminenspital Der Stat Wien, Vienna
Jayesh Mehta, Northwestern University, Chicago, USA
Giampaolo Merlini, University of Pavia, Pavia, Italy
Philippe Moreau, Nantes, France
Gareth Morgan, Royal Marsden Hospital, London, United Kingdom
Antonio Palumbo, Cathedra Ematologia, Torino, Italy
Santiago Pavlovsky, Fundaleu, Buenos Aires, Argentina
Amara Nouel, Bolivar, Venezuela
Susie Novis, IMF, Los Angeles, California, USA
Raymond Powles, Leukaemia & Myeloma, Wimbledon, United Kingdom
Linda Pilarski, The University of Alberta, Edmonton Alberta, Canada
S Vincent Rajkumar, Mayo Clinic, Rochester, Minnesota, USA
Tony Reiman, Cross Cancer Institute, Canada
Paul Richardson, Dana Farber Cancer Institute, Boston, Massachusetts, USA
Angelina Rodriquez Morales, Bonco MetroPolitano de Sangre, Caracas, Venezuela
Kazuyuki Shimizu, Nagoya City Midori General Hospital, Nagoya, Japan
David Siegel, Hackensack, Cancer Center, Hackensack, New Jersey, USA
Guido Tricot, MIRT UAMS, Little Rock, Arkansas, USA
Jesus San Miguel, University of Salamanca, Salamanca, Spain
Seema Singhal, Northwestern University, Chicago, Illinois, USA
Pieter Sonneveld, Erasmus University Hospital, Rotterdam, The Netherlands
Keith Stewart, Mayo Clinic Scottsdale, Scottsdale, Arizona, USA
Patrizia Tosi, Bologna, Italy
Ingemar Turesson, Malmo, Sweden
Ivan Van Riet, Brussels Vrija University
David Vesole, St Vincent's Comprehensive Cancer Center, New York, USA
Donna Weber, MD Anderson, Houston, Texas, USA
Jan Westin, University of Gothenberg, Sweden
Keith Wheatley, University of Birmingham, Edgbaston, Birmingham, UK
Brian Van Ness, University of Minnesota, Minnesota, USA