International Myeloma Workshop Consensus Panel 1. Consensus recommendations for the uniform reporting of clinical trials: report of the International Myeloma Workshop Consensus Panel 1

This is the author’s manuscript

Original Citation:

Availability:
This version is available http://hdl.handle.net/2318/100273 since

Terms of use:
Open Access
Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)
**Abstract**

It is essential that there be consistency in the conduct, analysis, and reporting of clinical trial results in myeloma. The goal of the International Myeloma Workshop Consensus Panel 1 was to develop a set of guidelines for the uniform reporting of clinical trial results in myeloma. This paper provides a summary of the current response criteria in myeloma, detailed definitions for patient populations, lines of therapy, and specific endpoints. We propose that future clinical trials in myeloma follow the guidelines for reporting results proposed in this manuscript.

**Introduction**

The treatment of myeloma has evolved rapidly in the last decade.\(^1\) The introduction of several active new drugs and novel targeted investigational agents has resulted in numerous active clinical trials in every stage of the disease. Studies are being conducted worldwide, including an increasing number of multicenter, international trials.\(^2,3\) It is essential that there be consistency in the conduct, analysis, and reporting of clinical trial results. Unless uniform reporting requirements are adhered to, it will be impossible to compare results across trials or to accurately determine whether reported results are valid and reliable. The goal of the International Myeloma Workshop Consensus Panel 1 was to develop a set of guidelines for the uniform reporting of clinical trial results in myeloma. We recognize that some compromises have to be made to ensure that this guidance meets requirements that are practical in most countries, academic and community practices, and various groups conducting clinical trials in myeloma. We propose that future clinical trials in myeloma follow the guidelines proposed in this manuscript.

**Lines of therapy**

A line of therapy is defined as one or more cycles of a planned treatment program.\(^4\) This may consist of one or more planned cycles of single-agent therapy or combination therapy, as well as a sequence of treatments administered in a planned manner. For example, a planned treatment approach of induction therapy followed by autologous stem cell transplantation, followed by maintenance is considered one line of therapy. A new line of therapy starts when a planned course of therapy is modified to include other treatment agents (alone or in combination) as a result of disease progression, relapse, or toxicity. A new line of therapy also starts when a planned period of observation off therapy is interrupted by a need for additional treatment for the disease.

**Definition of patient populations**
The terms used to define patient populations studied should be standardized. The terms “relapsed,” and “refractory,” when used to describe patient populations tested in clinical trials, should adhere to the definitions listed in this section. These definitions are based on a recent American Society of Hematology–Food and Drug Administration panel on endpoints in myeloma. We also propose that, when new clinical trials are initiated, these definitions be used in eligibility criteria to ensure uniformity across trials.

**Refractory myeloma**

Refractory myeloma is defined as disease that is nonresponsive while on primary or salvage therapy, or progresses within 60 days of last therapy. Nonresponsive disease is defined as either failure to achieve minimal response or development of progressive disease (PD) while on therapy. There are 2 categories of refractory myeloma: “relapsed-and-refractory myeloma” and “primary refractory myeloma.”

**Relapsed and refractory myeloma.**

Relapsed and refractory myeloma is defined as disease that is nonresponsive while on salvage therapy, or progresses within 60 days of last therapy in patients who have achieved minimal response (MR) or better at some point previously before then progressing in their disease course.

**Primary refractory myeloma.**

Primary refractory myeloma is defined as disease that is nonresponsive in patients who have never achieved a minimal response or better with any therapy. It includes patients who never achieve MR or better in whom there is no significant change in M protein and no evidence of clinical progression as well as primary refractory, PD where patients meet criteria for true PD. On reporting treatment efficacy for primary refractory patients, the efficacy in these 2 subgroups (“nonresponding-nonprogressive” and “progressive”) should be separately specified.

**Relapsed myeloma**

Relapsed myeloma is defined as previously treated myeloma that progresses and requires the initiation of salvage therapy but does not meet criteria for either “primary refractory myeloma” or “relapsed-and-refractory myeloma” categories.

**Additional qualifiers**

When possible, if a clinical trial is targeted to a specific population, it would be best to provide additional qualifiers that describe more precisely the population being studied, for example, “relapsed and refractory to immunomodulatory therapy” or “relapsed and refractory to bortezomib.” Prognostic factors, such as stage and cytogenetic information, should be considered as stratification factors at trial entry.

**Response criteria**

The International Myeloma Working Group (IMWG) uniform response criteria should be used in future clinical trials, with additional clarifications as listed in this section. The IMWG uniform response criteria were developed from the European Group for Blood and Bone Marrow Transplant/International Bone Marrow Transplant Registry/American Bone Marrow Transplant Registry published criteria, commonly referred to as the Blade criteria or the European Group for Blood and Bone Marrow Transplant criteria, with revisions and improvements that aid uniform reporting. These include the addition of free light chain (FLC) response and progression criteria for patients without measurable disease, modification of the
definition for disease progression for patients in complete response (CR), and addition of very good partial response (VGPR) and stringent response categories.

The panel endorsed the definitions of partial response (PR), VGPR, CR, PD, and stable disease according to IMWG. Of note, there was unanimous consensus that PD for patients in CR should be defined as per the IMWG criteria. CR patients will need to progress to the same level as VGPR and PR patients to be considered PD. A positive immunofixation alone is therefore not sufficient.\(^9,^{10}\)

The need for bone marrow confirmation of CR was discussed in detail, but new data showed that up to 14% of patients with immunofixation-negative CR may have more than or equal to 5% plasma cells in the marrow.\(^{11}\) Bone marrow confirmation is required for coding CR, and the panel recommends no change to the CR definition in this regard.

The clarifications and additions to the IMWG criteria discussed in this section were recommended and approved by the panel. The IMWG criteria for response and progression incorporating published errata and clarifications,\(^7,^{12,13}\) updated definition of stringent CR, and additional clarifications are listed in Tables 1 and 2.
The panel approved a definition of immunophenotypic CR to be incorporated into the IMWG criteria (Table 2). This requires absence of phenotypically aberrant plasma cells (clonal) in bone marrow with a minimum of 1 million total bone marrow cells analyzed by multiparametric flow cytometry (with ≥ 4 colors).¹⁴
The panel approved a definition of molecular CR to be incorporated into the IMWG criteria. Molecular CR is defined as CR plus negative allele-specific oligonucleotide polymerase chain reaction (sensitivity $10^{-5}$; Table 2).

Minimal response

The panel concurred with a recent American Society of Hematology-Food and Drug Administration panel\(^5\) that, for patients with relapsed and/or refractory myeloma, MR should be reported separately in clinical trials (Table 2). When MR is reported, the specific rate of MR should be distinguished from PR or better to make clinical trial comparisons possible.

Additional important clarifications

The following clarifications to IMWG criteria were made for coding CR in patients in whom the only measurable disease is by serum FLC levels (Table 1). In these patients, CR requires negative serum and urine immunofixation plus a normal FLC ratio of 0.26 to 1.65, on 2 consecutive assessments. Similarly, to code VGPR in such patients, a more than 90% decrease in the difference between involved and uninvolved FLC levels is required on 2 consecutive assessments. These were inadvertently omitted from the IMWG criteria.\(^12\) Some laboratories may have a slightly different reference range for the FLC ratio than 0.26 to 1.65. In these situations, it is appropriate to define normal FLC ratio using those used in the given laboratory.

Second, the panel clarified that bone marrow criteria for PD are to be used only in patients without “measurable disease” as defined in the IMWG criteria\(^7\) by M protein and by FLC levels. The “lowest response value” in determining the nadir for PD assessment does not need to be a confirmed value.

Third, the panel recommended that, if a patient has more than one M protein spike in the serum (or urine), the M protein to be followed for assessing response is only the one that meets IMWG criteria for “measurable” M protein level IMWG criteria.\(^7\) If more than one M protein spikes meet the criteria for measurable disease, then both need to be followed for response.

Fourth, the panel agreed that magnetic resonance imaging and positron emission tomography-computed tomography findings will not be incorporated formally into the response criteria for purposes of assessing
depth of response, but additional single-center studies are encouraged. Further validation of new aspects of the IMWG criteria will also be needed as agreed at the recent American Society of Hematology-Food and Drug Administration panel.

Finally, it is recommended that the time at which response assessment was conducted should be reported. In addition, the time to best response should also be reported.

**Reporting of efficacy results**

All efficacy results for primary endpoints should be reported only on an intent-to-treat basis. In the case of secondary endpoints, in addition to intent-to-treat results, results based on actual treatment received can also be reported. The reporting of results in subsets of patients restricted to those who completed certain duration of therapy should be avoided. All patients who were registered and met eligibility criteria regardless of whether they actually received therapy for a meaningful period (or not at all) should be in the denominator for all efficacy calculations. Response assessments should be performed before the next therapy is initiated.

In all clinical trials, patients should be followed every 1 to 2 months until PD to enable accurate calculation of time to progression (TTP) and progression-free survival (PFS).

**Essential efficacy measures in phase 3 trials**

Regardless of the primary endpoint studies, all phase 3 studies should report overall survival, TTP, PFS, duration of response (DOR), and if possible, time to next treatment (TNT), 5-year overall survival rate, and 10-year overall survival rate. The definitions of TTP, PFS, and DOR are listed in Table 3. It is particularly important that both TTP and PFS be reported. Where possible, details of any crossover should be provided.

### Table 3: Definitions of time to event endpoints

<table>
<thead>
<tr>
<th>Endpoint</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTP</td>
<td>Duration from start of treatment to disease progression, with deaths from causes other than progression counted.</td>
</tr>
<tr>
<td>PFS</td>
<td>Duration from start of treatment to disease progression or death (regardless of cause of death), whichever comes first.</td>
</tr>
<tr>
<td>EFS</td>
<td>The definition of EFS depends on how “event” is defined. In many studies, the definition of EFS used is the same as PFS. EFS may include additional “events” that are considered to be of importance because death and progression, including serious drug toxicity.</td>
</tr>
<tr>
<td>DFS</td>
<td>Duration from the start of CR to the first relapse or death. EFS applies only to patients in complete remission.</td>
</tr>
<tr>
<td>DOR</td>
<td>Duration from first observation of PR to the time of disease progression or death from causes other than progression.</td>
</tr>
</tbody>
</table>

TNT

TNT is difficult to accurately compare, except in double-blind studies, but it is clearly important to report TNT in future phase 3 trials. TNT is defined time from registration on trial to next treatment or death of any cause, whichever comes first. To accurately define TNT, next treatment should start uniformly in clinical
practice. The consensus is that the next treatment should start when there is either clinical relapse or a significant paraprotein relapse.

Clinical relapse is defined using the definition of clinical relapse in the IMWG criteria.\(^7\) In the IMWG criteria, clinical relapse is defined as requiring one or more of the following direct indicators of increasing disease and/or end-organ dysfunction that are considered related to the underlying plasma cell proliferative disorder:

1. Development of new soft tissue plasmacytomas or bone lesions on skeletal survey, magnetic resonance imaging, or other imaging
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
3. Hypercalcemia (> 11.5 mg/dL; > 2.875mM/L)
4. Decrease in hemoglobin of more than 2 g/dL (1.25mM) or to less than 10 g/dL
5. Rise in serum creatinine by more than or equal to 2 mg/dL (≥ 177mM/L)
6. Hyperviscosity

In some patients, bone pain may be the initial symptom of relapse in the absence of any of the features listed in “TNT.” However, bone pain without imaging confirmation is not adequate to meet these criteria in trials.

In patients who do not have clinical relapse, a significant paraprotein relapse is defined as doubling of the M-component in 2 consecutive measurements separated by less than or equal to 2 months; or an increase in the absolute levels of serum M protein by more than or equal to 1 g/dL, or urine M protein by more than or equal to 500 mg/24 hours, or involved FLC level by more than or equal to 20 mg/dL (plus an abnormal FLC ratio) in 2 consecutive measurements separated by less than or equal to 2 months. This definition of “paraprotein relapse” represents the rate of rise or absolute level of increase in M protein at which the panel considered that myeloma therapy should be restarted in relapsing patients in clinical practice, even if signs and symptoms of new end-organ damage are not yet apparent.

Summary and future directions

This paper summarizes, clarifies, and updates current response criteria in myeloma. We have provided detailed definitions for patient populations, lines of therapy, and specific endpoints. We propose that future clinical trials in myeloma follow the guidelines for monitoring patients and reporting results proposed in this manuscript. These criteria will most probably change with time as the technology improves and more sensitive tests become available. We also need to develop criteria to assess the efficacy of therapy for earlier stages of the disease, such as smoldering multiple myeloma given the interest in preventive clinical trials. Finally, we need to quickly develop and validate response criteria that incorporate gene expression profiling and imaging techniques, such as positron emission tomography.

References


