Approach to the treatment of multiple myeloma: a clash of philosophies

S. Vincent Rajkumar, Gösta Gahrton, and P. Leif Bergsagel

Division of Hematology, Mayo Clinic, Rochester, MN; Karolinska Institutet, Huddinge, Stockholm, Sweden; and Division of Hematology and Oncology, Mayo Clinic, Scottsdale, AZ

In this Perspective, we summarize some of the most contentious issues surrounding diagnosis and treatment of myeloma. We outline how a fundamental clash of philosophies, cure versus control, may be at the heart of many of the controversies. From the very definition of the disease to risk stratification to the validity of current clinical trial endpoints, we highlight the major areas of debate and provide alternative viewpoints that have implications for trial design and interpretation, as well as clinical practice. (Blood. 2011;118(12):3205-3211)

Introduction

Multiple myeloma (MM) is a malignant plasma cell disorder that accounts for ~10% of all hematologic malignancies. The diagnosis of MM requires 10% or more clonal plasma cells in the bone marrow examination or the presence of a biopsy-proven plasmacytoma, plus evidence of end-organ damage (eg, anemia, hypercalcemia, lytic bone lesions, or renal failure) attributable to the underlying plasma cell disorder.3

The current treatment of myeloma varies considerably across institutions and across countries. Part of this variation is the result of availability and cost of new drugs, but part of it may be related to differing treatment philosophies and alternative interpretations of the available data. In general, initial therapy is decided based on eligibility for autologous stem cell transplantation (ASCT). Thus, patients eligible for ASCT are treated with a non–melphalan-containing induction regimen followed by transplantation (early transplantation approach), although some opt for a strategy of continuing initial therapy reserving ASCT for first relapse (delayed transplantation approach).1 In randomized trials so far, both the early and delayed ASCT approaches appear to have the same overall survival (OS).4,5 Patients who are not eligible for ASCT have traditionally been treated with a melphalan-containing regimen for 12-18 months, but some of the regimens used in ASCT candidates are being used in this patient population as well. The specific regimen chosen for initial therapy varies widely, and the use of consolidation and/or maintenance after initial therapy are still unsettled. Most patients inevitably relapse. Although the options for the treatment of relapsed disease have increased, the remission duration decreases with each salvage regimen.6

The overall prognosis of MM depends on a variety of host- and disease-related factors summarized in Table 1.7 However, there is no question that the survival of MM patients has improved significantly in the last decade.1,2 This improvement can be attributed to active new agents thalidomide,8 bortezomib,9,10 and lenalidomide11,12; optimal use of ASCT13-15; and improvements in supportive care.16 The role of allogeneic transplantation is controversial but may have contributed to improved outcome for some patients as well.17,18 And the future is promising, with many other potentially active agents in the pipeline.19,20 However, there are certain fundamental issues that investigators in the field disagree on, which in our opinion need to be carefully studied and debated.

Disease definition

There are 2 critical problems with the very definition of the disease that merit attention. First, unlike most other malignancies, the diagnosis of symptomatic MM requiring treatment is made by a clinician, based on the presence or absence of clinical findings: hypercalcemia, renal failure, anemia, and lytic bone lesions, which are considered attributable to the clonal plasma cell proliferative disorder.3 There is considerable subjectivity involved in interpreting which patient meets the clinical requirement because the attribution is often subjective and the definition does not call for any specific laboratory proof for the attribution. As tests to detect bone disease become increasingly sensitive, the situation is ripe for patients who in years past would have been considered to have smoldering MM increasingly being classified as symptomatic MM. It should be recognized that the availability of effective new therapies, the increased “education” of patients that goes with it, and the use of sensitive tests to detect disease (whole body positron emission tomography and magnetic resonance imaging) may result in stage migration. This can then lead to treatment outcomes that appear better than they actually are, or worse, result in the treatment of patients who do not need it.

Second, MM is defined as a single disease entity based on its phenotype.24,25 If we could turn back time, we may not call it one disease. There are many cytogenetically (and to some extent clinically) distinct diseases that are currently lumped together and referred to as MM. Although, as in other hematologic malignancies, it is possible that the benefits and risks of the various treatment approaches may not be universally applicable to all cytogenetic categories, this is not aligned with commercial reality.

We think that additional work needs to be urgently done to revise the disease definition. In the short-term, the International Myeloma Working Group is working on collecting the necessary data to update the disease definition. In the long-term, we need to
be in line with other malignancies and identify pathologic and molecular markers that differentiate benign clonal expansion from true malignancy. Until this happens, we need to recognize that comparison of recent trial results to historical outcomes may not be valid because the intensity of the diagnostic workup and the willingness to treat have changed over time, but the disease definition has not. Importantly, we also need to consider collaborative trials for individual cytogenetic categories.

**Risk-adapted approach to treatment**

Survival for an individual patient with MM varies according to age, stage, and performance status, as in any other malignancy. Appropriately, therapy is usually adjusted based on these factors. However, the majority of patients are being approached with a uniform philosophical approach: achieve and maintain complete response (CR). This needs to change. The prognosis of MM varies considerably between cytogenetic categories and even within each cytogenetic entity. For example, the presence of high-risk gene-expression profiling signature and 17p deletion identifies high-risk subsets within each nonoverlapping cytogenetic category of MM. It is not reasonable, in our opinion, to approach patients with dramatically different survival probabilities with an identical therapeutic philosophy. Even if the outcomes with various treatment strategies are similar, a patient’s outlook on life and tolerance for risk can be affected greatly by expected survival.

More importantly, standard-risk patients who have expected median survivals in excess of 10 years when given full informed consent of the treatment options, may be reluctant to choose regimens that carry an early risk of irreversible toxicity. For such patients, even the route of administration and the number of monthly visits to a healthcare provider may be factors in choosing a regimen.

Some investigators think that most treatments work best in the standard-risk population; therefore, one should use combinations that have the highest probability of achieving CR for this patient population. Although surrogate endpoints, such as CR or progression-free survival (PFS), can be improved (predictably) by increasing the number of active drugs that are given at the time of initial diagnosis, there are little data that OS or quality of life is improved as long as all patients have access to the same drugs. In general, only a subset of patients randomized to the control arms of myeloma studies that show an advantage in OS have had access to the experimental drug at relapse. Although a trial design that lacks a built-in crossover has been very important for the timely regulatory approval of new drugs, it is not useful for the clinical question of sequential versus combination use of agents with known activity, or with regards to timing of use in the overall treatment strategy. Distressingly, most myeloma phase 3 trials are designed with surrogate endpoints and typically fail to have well-designed quality-of-life studies built in that can shed light on how patients in different risk categories benefit from various interventions.

These issues are magnified by the fact that patients often have little choice in selection of treatment agents, and worse, do not have the data required to make informed decisions. Thus, treatment is usually based on “institutional practice” or “flavor of the month.” As discussed in “The CR myth: there is no free lunch” and “PFS,” if regimens are judged by suboptimal surrogate endpoints and if the distinction between surrogate endpoints versus real endpoints is blurred, it makes it easy for the selection of “flavor of the month” to be unduly influenced by early phase (phase 1, 2) studies, as well as by a variety of other irrelevant factors. Unfortunately, it is impossible to say whether an individual patient has benefited by achieving a surrogate endpoint, whereas it is often easy to see how they have been harmed (eg, irreversible neuropathy, pulmonary embolus, second primary malignancy).

**The CR myth: there is no free lunch**

Several papers have emphasized the importance of CR in MM.26,27,31-35 This is a difficult topic to cast doubt on because, after all, who can be against CR? The main reason to be cautious in espousing CR is the correlation between increased response rates and increased toxicity. Striving blindly for CR may lead to unacceptable and unnecessary toxicity for some patients and come at too great a price. However, oncologists willing to say that CR is not an important endpoint can quickly attain the pariah status. Despite this, we wish to highlight some important caveats about CR and illustrate its limited use as a surrogate endpoint in a disease, such as MM.

First, it is well known that responders will always appear to do better than nonresponders, regardless of whether or not the treatment works.21 Nevertheless, the habit of comparing survival between those who achieve CR versus those who do not has not diminished. The fact that this type of analysis is fundamentally flawed needs to be recognized; responders should not be directly compared with nonresponders, and they will always appear to do better regardless of whether the treatment works or not.21 Further, attainment of CR is as much a function of tumor biology as it is a reflection of a therapy’s value. To truly study the value of CR, we need either randomized trials comparing 2 strategies: one...
in which CR is a goal of therapy and treatments are modified as needed to achieve that goal versus another in which a specified therapy is given regardless and not modified based on CR status; or at least appropriate landmark analyses at various time intervals.36

Second, unlike curable cancers, such as large cell lymphoma, CR in MM is not a surrogate for cure but rather a marker of profound tumor reduction. It is not a true CR because most CR patients will relapse with time. Indeed, in most studies that proclaim the benefit of CR, the median CR duration is short.

Third, in most cases, it is probably impossible (with the possible exception of allogeneic transplantation where true complete eradication is possible for a minority of patients), unnecessary, and prohibitively toxic to attempt to eradicate all clonal plasma cells. Unlike other cancers where benign clonal expansion can be distinguished histopathologically from malignant cells, such discrimination is not currently possible in MM.1 Thus, achieving immunofixation-negative CR status may need eradication of not just MM cells but also any premalignant monoclonal plasma cell populations present in a given patient. The existence of such clones is not in doubt and is evidenced by the fact that premalignant populations present in a given patient. The existence of such clones just MM cells but also any premalignant monoclonal plasma cell immunofixation-negative CR status may need eradication of not every MM patient needs or benefits from achieving CR.36 These data compare outcomes between CR vs no CR) that not every

The discussion on disease definition, risk stratification, and CR is very important, it pales in comparison with the interpretation of PFS in MM trials. Indeed, nothing perturbs us more than that the indiscriminate adoption of PFS as a major endpoint without regard to the specific clinical question in hand. To be clear, we fully support the use of PFS (or the related endpoint of time to progression [TTP]) as an indicator of drug activity, and as an appropriate endpoint in many instances.44 For example, PFS is an appropriate endpoint when determining whether or not a particular drug (or treatment modality, such as autologous or allogeneic transplantation) has efficacy. Similarly, PFS is indeed the appropriate endpoint for the approval of new drugs in MM. This is how the new drugs, such as bortezomib, lenalidomide, and liposomal doxorubicin, received full approval.9,45-47 Indeed, without PFS as the primary endpoint, there will be a delay in getting important new drugs into practice. When the question is whether or not a new drug is of clinical benefit, PFS is an excellent surrogate for OS, and possibly patient-reported quality of life.

But the metrics needed for a specific modality to be considered potentially effective or for a new drug to receive Food and Drug Administration approval (proof of clinical benefit) differ completely from the metrics that need to be applied when evaluating nonregulatory trials of drugs already on the market, where the goal is to simply to determine the timing of an approved drug (early vs delayed use) in the overall treatment strategy (Figure 1A).

Why the distinction? In regulatory clinical trials, endpoints such as PFS and TTP are meaningful in a disease, such as MM, because a significant prolongation of those endpoints can be reasonably indicative of clinical benefit.9,45-47 Because the drug being tested is not approved, patients in the control arm (and patients at large) do not have the option of getting the drug later on in the disease course. Thus, the question being asked is whether or not the drug should be available for MM patients.

In contrast, in most nonregulatory MM trials, PFS and TTP do not carry the same value because patients in the control arm do have the option of getting the same drug later. In these studies, mostly designed to promote early or more prolonged use of an approved drug, the question is not whether or not the drug should be used, but rather when in the treatment course it is best used. Here, PFS (or TTP) improvement with early use does not necessarily imply clinical benefit (ie, prolonged OS or patient-reported quality of life). To show that the drug is better given early in the treatment course, PFS (or TTP) in the control arm must ideally be measured after the patient has failed the treatment in question when administered at first relapse (Figure 1B). In essence, when evaluating the 2 approaches (early vs late treatment), one must compare

### Table 2. Results of 2 recent randomized trials showing lack of correlation between regimens that provide better depth of response and survival

<table>
<thead>
<tr>
<th>Reference</th>
<th>Regimen</th>
<th>No. of patients</th>
<th>CR, %</th>
<th>CR plus VGPR, %</th>
<th>Median PFS, months</th>
<th>P for PFS</th>
<th>OS</th>
<th>P for OS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harousseau et al40 (2010)</td>
<td>VAD</td>
<td>242</td>
<td>9</td>
<td>37</td>
<td>30</td>
<td>.06</td>
<td>3-yr, 77%</td>
<td>.5</td>
</tr>
<tr>
<td></td>
<td>TD</td>
<td>240</td>
<td>16</td>
<td>54</td>
<td>36</td>
<td>.36</td>
<td>3-yr, 81%</td>
<td></td>
</tr>
<tr>
<td>Ludwig et al41 (2009)</td>
<td>MP</td>
<td>143</td>
<td>2</td>
<td>13</td>
<td>21</td>
<td>.1</td>
<td>Median, 49 mo</td>
<td>.02*</td>
</tr>
<tr>
<td></td>
<td>TD</td>
<td>145</td>
<td>2</td>
<td>26</td>
<td>17</td>
<td></td>
<td>Median, 42 mo</td>
<td></td>
</tr>
</tbody>
</table>

VGPR indicates very good partial response; VD, bortezomib plus dexamethasone; VAD, vincristine, doxorubicin, dexamethasone; MP, melphalan plus prednisone; and TD, thalidomide plus dexamethasone.

*In favor of MP over TD.
the time to approach failure, which in the case of early treatment is the initial PFS/TTP, but for the late treatment is the time from randomization until relapse from the late treatment (second relapse).

The concepts discussed in the preceding paragraph can be best illustrated by some examples. If administering lenalidomide in MM shows improved PFS over not giving lenalidomide, then the drug has clinical benefit and should be made available to patients. This is of course what happened with the regulatory trials that led to the approval of the drug (Figure 2A).45,46 On the other hand, consider trials testing the role of lenalidomide as maintenance therapy for patients with MM after ASCT (Figure 2B).48,49 In this scenario, prolongation of PFS in patients receiving lenalidomide is not a surprising finding. After all, the drug is known to be active. It also does not imply clinical benefit. Patients in the control arm can always take lenalidomide at progression and may derive that same PFS benefit later.

Unless OS or validated patient-reported outcomes are clearly superior, the same argument also holds true for trials that use PFS as the primary endpoint to compare 3 known active agents versus 2 in the treatment of newly diagnosed MM: patients who received only 2 drugs as initial therapy can always take (or add) the third drug at the time of first relapse.

One may question what the big deal is (Figure 3). Why not go with treatments that prolong PFS anyway in the nonregulatory settings discussed in the preceding 2 paragraphs? Is there not a potential for the drug to actually prolong OS when given early in the disease course before the development of additional cytogenetic abnormalities or microenvironmental independence? There are at least 5 reasons why we need to be cautious in this regard, besides the compelling arguments by Saltz that the very term PFS is erroneous and should be changed to something more meaningful, such as progression-free interval.23

First, unlike a disease, such as large cell lymphoma, where prolonged PFS (or disease free survival) is a surrogate for the cure rate of patients, PFS in MM means something different: the timing at which the next drug (or regimen) needs to be introduced into the treatment pathway. Prolongation of OS in MM is almost always associated with prolonged PFS, but not the other way around.40,41,50,51 Treatment of MM is like a marathon, rather than a sprint. It is for most patients (ie, non–high-risk MM) a chronic malignancy, in which patients relapse repeatedly, and using all active drugs early on may be the equivalent of a marathon runner running the first mile at breakneck speed. He/she may finish the race at the same time as a runner who spreads out the energy a little better, but the remaining 25 miles would have been a lot more pleasant. This in fact is what is happening in most recent trials, in which PFS is prolonged with more aggressive initial therapy, but OS is identical. In this situation, without true quality of life studies, it is impossible to determine clinical benefit.

Second, the use of PFS as a surrogate endpoint to justify earlier and more prolonged use of new treatments results in more people taking treatment (or an extra drug/intervention) when they do not need it. Delayed use can achieve the same clinical benefit (similar OS) despite fewer patients requiring the given therapy. For example, in one randomized trial, delayed ASCT was needed in
only 35% of the randomized population to get the same benefit as early SCT in terms of OS. In other words, patients in the control arm need to receive a smaller PFS benefit to equalize as shown in the case of the lenalidomide maintenance setting in Figure 2B. For example, the routine use of lenalidomide maintenance in the United States in our estimate will cost $1 billion, and for such a price we need to understand exactly how much OS benefit we get in return. CR, TTP, and PFS are simply means to an end. Not the end itself.

Third, when there is no proof that OS is prolonged by a clinically significant duration, and all we have is improvement in PFS, we need to consider the risks (sometimes unforeseen) of such treatments carefully. This issue has been brought to the forefront by the thorny issue of second cancers with lenalidomide maintenance. There are now emerging data\textsuperscript{48,49} and a Food and Drug Administration warning that the use of lenalidomide as maintenance may increase the risk of second cancers. Patients are being sent letters that this risk may exist, prompting worried phone calls. Clearly, if OS is prolonged with such therapy, we would be much less concerned because it factors in the excess second cancer risk. But prolongation of PFS alone is not adequate because the seriousness and implications of a second cancer are quite different from a rise in the monoclonal (M) protein by 25% (which is all that it takes to qualify as “progression” in MM). Indeed, we know of no patient who would trade a small M protein increase from 1.0–1.5 g/dL for a second cancer. In these circumstances, we need to be cautious. We need to consider whether all patients benefit equally; whether one subgroup may benefit, whereas another may not; whether some patients may actually be harmed. Thankfully, in the case of lenalidomide maintenance, an OS advantage is emerging in one of 2 randomized clinical trials, and the risk of second cancers may be overcome if a significant survival benefit is confirmed.

Fourth, the oft-repeated argument that OS is not a good endpoint in MM because of the availability of numerous salvage options is flawed and perpetuates a myth and a misunderstanding. Indeed, this is precisely the reason one needs to look for OS improvement rather than PFS improvement in this disease. If survival is the same eventually, what exactly is the point of improved PFS by giving an additional new drug early in the disease course? Is there not an increased risk of toxicity when more drugs are combined? Is there not a risk that patients are subject to potentially irreversible risks earlier in their life?

Finally, some have argued that OS differences are difficult to achieve and require large unrealistic sample sizes. True. But just because something is easy or difficult should not be a reason to justify adoption of suboptimal endpoints to change practice. We do not have an issue with MM trials that are designed with PFS as the endpoint from purely a feasibility standpoint. Our quarrel is with the interpretation. If such trials show PFS prolongation, we would rather interpret it as an interesting hypothesis that then needs to be tested for presence or absence of a true clinical benefit.

Cure versus control

The points discussed in this Perspective are indicative of a broader clash of philosophies in MM: cure versus control. Those favoring the curative approach think that incorporation of all active agents early in the disease course targets the disease at its most vulnerable
stage before the development of microenvironmental independence and secondary cytogenetic abnormalities. Those favoring the disease-control approach prefer a risk-adapted approach in which an aggressive strategy is pursued in high-risk patients (17p deletion, high-risk gene-expression profiling signature, t4;14, t14;16, t14;20), but most others are offered an alternative sequential therapy approach that minimizes toxicity, maximizes quality of life, and emphasizes patient preference. The cure-versus-control debate colors our interpretation of clinical trial results, with well-meaning investigators interpreting the same clinical trial data in opposite ways, depending on whether they ascribe to the cure or control philosophy as they approach the care of patients with MM.

In clinical practice, we prefer the control approach for most patients, acknowledging that high-risk subgroups need a more aggressive approach targeting CR or cure. For example, younger patients may opt for early intensive strategies, such as allogeneic transplantation, accepting high treatment-related mortality in exchange for a chance at long-term survival. In a recent editorial, Gatenby argued that trying to control cancer may be a better strategy in some circumstances than striving to cure it. Although early use of multidrug combination strategies and maintenance therapy improve CR and PFS, we need to strive for clear evidence of OS improvement to change practice.

In clinical trials, we support well-designed investigations of both the curative and control approaches. We do see the merits and arguments of the curative approach, and we recommend head-to-head trials comparing the 2 strategies. We hope that, as our treatment armamentarium increases over time, we will be able to achieve ever higher response rates with ever lower toxicity and the pendulum will gradually shift from control to cure.

Acknowledgments

This work was supported in part by the National Cancer Institute (grants CA107476, CA62242, CA1000707, and CA83724).

Authorship

Contribution: S.V.R., G.G., and P.L.B. researched, provided critical review, wrote the paper, and reviewed and approved the final manuscript.

Conflict-of-interest disclosure: G.G. is on the advisory board for Fujimoto Pharmaceutical Company and Karolinska Innovation and has received honoraria from Celgene. P.L.B. has participated in advisory boards for Novartis, Celgene, and Chiron and receives research support for a clinical trial from Genentech. The remaining author declares no competing financial interests.

Correspondence: S. Vincent Rajkumar, Division of Hematology, Mayo Clinic College of Medicine, Mayo Clinic, 200 First St SW, Rochester, MN 55905; e-mail: rajkumar.vincent@mayo.edu.

References

TREATMENT OF MULTIPLE MYELOMA


Approach to the treatment of multiple myeloma: a clash of philosophies

S. Vincent Rajkumar, Gösta Gahrton and P. Leif Bergsagel