Superior outcomes associated with complete response in newly diagnosed multiple myeloma patients treated with nonintensive therapy: analysis of the phase 3 VISTA study of bortezomib plus melphalan-prednisone versus melphalan-prednisone

Jean-Luc Harousseau, Antonio Palumbo, Paul G. Richardson, Rudolf Schlag, Meletios A. Dimopoulos, Ofer Shpilberg, Martin Kropff, Alain Kentos, Michele Cavo, Anatoly Golenkov, Mieczyslaw Komarnicki, Maria-Victoria Mateos, Dixie-Lee Esseltine, Andrew Cakana, Kevin Liu, William Deraedt, Helgi van de Velde, Jesús F. San Miguel

Abstract

The phase 3 Velcade as Initial Standard Therapy in Multiple Myeloma: Assessment with Melphalan and Prednisone study in newly diagnosed multiple myeloma patients ineligible for high-dose therapy demonstrated that bortezomib-melphalan-prednisone (VMP) was superior to melphalan-prednisone across all efficacy end points. We assessed the prognostic impact of response on time-to-event parameters in the intent-to-treat population. Patients received nine 6-week cycles of treatment. Time to progression, time to next therapy, and treatment-free interval were associated with quality of response. When European Group for Blood and Marrow Transplantation criteria were used, complete response (CR) was associated with significantly longer time to progression (hazard ratio [HR] = 0.45, \( P = .004 \)), time to next therapy (HR = 0.46, \( P = .0004 \)), and treatment-free interval (HR = 0.38, \( P < .0001 \)) versus partial response, but there was no significant difference in overall survival (HR = 0.87, \( P = .54 \)); similar differences were seen with CR versus very good partial response by uniform criteria. Quality of response improved with prolonged VMP treatment, with 28% of CRs achieved during cycles 5-9. CR duration appeared similar among patients with “early” (cycles 1-4) and “late” CRs (cycles 5-9) and among patients receiving 9 versus < 9 cycles of bortezomib within VMP. These results highlight that CR is an important treatment goal and support prolonged VMP therapy to achieve maximal response. This study is registered at http://www.clinicaltrials.gov as NCT00111319.

Introduction

Multiple myeloma (MM) is an incurable plasma-cell neoplasm. The primary aim of current treatment strategy is to improve long-term outcomes, including progression-free survival (PFS) and overall survival (OS),\(^1\) as well as health-related quality of life (HRQoL), which is affected by treatment effect, toxicity, and change in disease status.\(^2\) The introduction of high-dose therapy and stem-cell transplant (HDT-SCT), which is now the standard of care for eligible patients ≤ 65 years of age, and the novel agents bortezomib, thalidomide, and lenalidomide has been associated with improved OS in the past 10 years.\(^3\)–\(^5\)

The authors of studies in previously untreated MM patients who received HDT-SCT have demonstrated that quality of response, notably achievement of complete response (CR) or at least very good partial response (VGPR), is prognostic for improved long-term outcomes.\(^6\)–\(^13\) However, in patients ineligible for HDT-SCT, there is limited evidence for such a correlation between quality of response and long-term outcomes.\(^14\) This may be because until recently, standard frontline therapies for patients ineligible for HDT-SCT did not result in substantial rates of CR or CR/VGPR, thereby preventing meaningful analysis of outcomes according to
response. Now, novel agent-based treatment options for previously untreated nontransplantation MM patients are demonstrating high CR and CR/VGPR rates, superior to those achieved with previous standards of care. The use of bortezomib in combination with conventional therapies has resulted in increased CR rates associated with improved outcomes.

The phase 3 Velcade as Initial Standard Therapy in Multiple Myeloma: Assessment with Melphalan and Prednisone (VISTA) study compared bortezomib plus melphalan-prednisone (VMP) versus melphalan-prednisone alone in previously untreated patients with MM who were ineligible for high-dose therapy. This study demonstrated that VMP was superior to MP across all response rate and time-to-event efficacy end points and resulted in a very high CR rate of 30%. Thus, we investigated whether, as seen in the transplant setting, achievement of CR was associated with improved outcome. Here, we assess the differential prognostic impact of best response on time-to-event parameters in VISTA in the substantial population of patients achieving CR and partial response (PR) and evaluate time-to-event distributions in the VMP arm by quality of response. In addition, we evaluate the impact of timing of CR on outcome in patients receiving VMP.

**Methods**

**Study design**

The VISTA study design has been described previously (http://www.clinicaltrials.gov, NCT00111319). In brief, previously untreated MM patients were randomized to receive 9 6-week cycles of VMP, comprising bortezomib 1.3 mg/m², days 1, 4, 8, 11, 22, 25, 29, and 32, cycles 1-4, and days 1, 8, 22, and 29, cycles 5-9; melphalan 9 mg/m², days 1-4, cycles 1-9; and prednisone 60 mg/m², days 1-4, cycles 1-9, or MP alone. The primary end point was time to progression (TTP). Response was determined by a prespecified computer algorithm and progression was determined by investigator assessment by use of the European Group for Blood and Marrow Transplantation (EBMT) criteria; a central laboratory was used for M-protein assessment. In a post hoc analysis, response was determined by use of the International Myeloma Working Group (IMWG) uniform criteria. As previously reported, the VISTA study was stopped after the third interim analysis on the basis of an independent recommendation from the data monitoring committee (data cutoff, June 15, 2007). Centralized collection of M-protein data was discontinued; therefore, subsequent data on response rates and TTP were not updated; however, data collection continued for subsequent therapy use and OS.

The VISTA study was approved by the Independent ethics committee/institutional review board at all participating institutions and was conducted in accordance with International Conference on Harmonization Guidelines on Good Clinical Practice and the principles of the Declaration of Helsinki. All patients provided written informed consent.

**Analysis of outcomes by response**

Patients who had measurable disease at baseline and who received at least one dose of study drug were evaluable for response. For the purposes of the present analysis, patients were classified by either the major response achieved (CR or PR), or as those with minimal or no response (< PR). By EBMT criteria, 102 (30%), 136 (40%), and 95 (28%) patients in the VMP arm and 12 (4%), 103 (31%), and 212 (64%) patients in the MP arm achieved CR, PR, and < PR, respectively. By IMWG criteria, 111 (33%), 28 (8%), 112 (33%), and 82 (24%) patients in the VMP arm and 13 (4%), 13 (4%), 102 (31%), and 199 (60%) in the MP arm achieved...
CR, VGPR, PR, and < PR, respectively.\textsuperscript{16} It was not possible to include the additional response category of near-CR\textsuperscript{20} in the present analysis because of the small number of patients achieving this response.

TTP, time to next therapy (TNT), treatment-free interval (TFI), and OS were evaluated in the VMP arm according to response by EBMT and IMWG criteria. Long-term outcomes (TTP, TNT, TFI, OS, and time from response/CR to subsequent therapy) also were evaluated in VMP patients achieving early CR, during treatment cycles 1-4, or late CR, during treatment cycles 5-9. Response and TTP data were from the third interim analysis (median follow-up was 16.3 months).\textsuperscript{16} TNT, TFI, and OS data were from the most recent updated analysis, with a data cutoff of March 16, 2009, and median follow-up was 36.7 months. For the same reasons as for response and TTP data, response and CR duration could not be updated from the third interim analysis; therefore, time from response/CR to subsequent therapy was analyzed as a surrogate for response/CR duration by the use of subsequent therapy data from the updated analysis.

**Statistical analysis**

Response duration was defined as the time from first evidence of response to start of subsequent therapy, and CR duration was the time from first evidence of CR to start of subsequent therapy, as surrogates for time to relapse, as described previously. TTP was defined as time from randomization to progression, TNT was time from randomization to next therapy, and TFI was length of time after therapy without treatment. Associations between TTP, TNT, TFI, and OS and response were examined in the intent-to-treat population, which included both VMP and MP patients, by multivariate Cox regression analysis with time-dependent covariates, adjusted for stratification factors. Covariates included age, sex, race, MM type, baseline Karnofsky performance status, and number of bone lesions; \( P \) values for comparisons of outcomes by response were determined by stratified log-rank tests. TNT, TTP, OS, TFI, and time from response/CR to subsequent therapy distributions according to best response achieved were estimated for the response-evaluable population in the VMP arm by Kaplan-Meier methodology.

**Results**

**Patients**

A total of 682 patients were randomized to receive VMP (\( n = 344 \)) or MP (\( n = 338 \)). The response-evaluable population included 337 and 331 patients in the VMP and MP arms, respectively. Patient demographics and baseline disease characteristics for the overall population have been reported previously\textsuperscript{16} and were typical of a transplant-ineligible patient population. Patient demographics and baseline characteristics in the response-evaluable population in the VMP arm according to response by EBMT criteria are listed in Table 1. Characteristics appeared generally similar between response subgroups; however, the proportion of patients aged \( \geq 75 \) years appeared greater in the < PR subgroup (42%) compared with the CR (27%) and PR (26%) subgroups. The proportion of patients who had a baseline \( \beta_2 \)-microglobulin level of \( > 5.5 \) mg/L appeared greater in those achieving CR versus PR or < PR (42% vs 26% or 32%). A greater proportion of patients achieving CR (42%) had ISS stage III disease, compared with patients achieving PR (29%) or < PR (33%). Among patients with ISS disease stage I who achieved CR, most were achieved early. In addition, the ratio of MM subtypes varied between response subgroups, with a lower proportion of immunoglobulin G MM and a greater proportion of immunoglobulin A MM patients within the CR subgroup compared with the PR or < PR subgroups.
Analyses of outcomes by response

Results of multivariate Cox regression analysis of outcomes according to response by EBMT criteria are shown for the intent-to-treat population (all patients in both the VMP and MP arms) in Table 2. CR was associated with significantly longer TTP (hazard ratio [HR] = 0.45, \( P = .004 \)), TNT (HR = 0.48, \( P < .0001 \)), and TFI (HR = 0.42, \( P < .0001 \)) versus PR. By contrast there was no significant difference between CR and PR in terms of OS (HR = 0.87, \( P = .54 \)). Furthermore, CR and PR were associated with significantly longer TTP, TNT, TFI, and OS versus < PR. Similar associations between response and outcomes were seen in patient subgroups who had ISS stage I/II or ISS stage III MM (supplemental Tables 1 and 2, available on the Blood Web site; see the Supplemental Materials link at the top of the online article).
Table 2

Table 2. Multivariate analysis of outcomes according to best response in the intent-to-treat population, in VMP and MP patients, by EBMT criteria

<table>
<thead>
<tr>
<th>Hazard ratio (95% CI)</th>
<th>Versus PR</th>
<th>Versus &lt; PR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TTP</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patients achieving CR</td>
<td>0.45 (0.26-0.70), P = .004</td>
<td>0.31 (0.19-0.56), P = .0001</td>
</tr>
<tr>
<td>Patients achieving PR</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td><strong>TNT</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patients achieving CR</td>
<td>0.49 (0.43-0.88), P = .0001</td>
<td>1.33 (0.82-0.97), P &lt; .0001</td>
</tr>
<tr>
<td>Patients achieving PR</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td><strong>TFI</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patients achieving CR</td>
<td>0.62 (0.44-0.89), P = .0005</td>
<td>—</td>
</tr>
<tr>
<td>Patients achieving PR</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td><strong>OS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patients achieving CR</td>
<td>0.97 (0.68-1.38), P = .54</td>
<td>0.32 (0.22-0.58), P = .0001</td>
</tr>
<tr>
<td>Patients achieving PR</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

CI indicates confidence interval; CR, complete response; EBMT, European Group for Blood and Marrow Transplantation; OS, overall survival; PR, partial response; TFI, treatment-free interval; TNT, time to next therapy; TTP, time to progression; VMP, bortezomib-melphalan-prednisone; MP, melphalan-prednisone; and —, not applicable.

Kaplan-Meier distributions of TTP, TNT, TFI, and OS in VMP patients achieving CR or PR by EBMT criteria reflect the findings of the multivariate analysis (Figure 1). Patients achieving CR had prolonged TTP, TFI, and TNT versus those achieving PR; OS appeared similar among patients achieving CR or PR to VMP although, as noted previously, by multivariate analysis, the HR was 0.87 in favor of CR.

Figure 1
Multivariate Cox regression analysis of outcomes according to response by IMWG criteria (Table 3) showed that CR was associated with significantly longer TTP (HR = 0.45, \( P = .019 \)), TNT (HR = 0.52, \( P = .004 \)), and TFI (HR = 0.45, \( P = .0008 \)) versus VGPR, whereas no significant differences were seen in outcomes between patients achieving VGPR and PR (Table 3). It should be noted that the number of patients achieving VGPR was small (n = 41). Analysis of outcomes in patients achieving CR versus VGPR + PR versus < PR by IMWG criteria demonstrated similar findings to the analysis of CR versus PR versus < PR by EBMT criteria (data not shown). In addition, CR + VGPR was associated with significantly longer TTP (HR = 0.60, \( P = .025 \)), TNT (HR = 0.65, \( P = .005 \)), and TFI (HR = 0.63, \( P = .002 \)) compared with PR, but there was no significant difference in OS (HR = 0.88, \( P = .54 \)). Kaplan-Meier distributions of TTP in patients achieving CR, CR + VGPR, VGPR, or PR by IMWG criteria with VMP demonstrated longer TTP (median not reached vs not reached vs 16.8 months vs 21.7 months) in patients who achieved CR or CR + VGPR (Figure 2).

Table 3

<table>
<thead>
<tr>
<th>Response</th>
<th>Hazard Ratio (95% CI)</th>
<th>( P ) Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR</td>
<td>0.45 (0.24-0.86), ( P = .019 )</td>
<td>0.0001</td>
</tr>
<tr>
<td>VGPR</td>
<td>0.50 (0.32-0.80), ( P = .013 )</td>
<td>0.0001</td>
</tr>
<tr>
<td>PR</td>
<td>0.59 (0.35-0.99), ( P = .025 )</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

CI indicates confidence interval; CR, complete response; IMWG, International Myeloma Working Group; OS, overall survival; PR, partial response; TFI, treatment-free interval; TNT, time to next therapy; TTP, time to progression; VGPR, very good partial response; and —, not applicable.
Impact of timing of CR on outcome

The median time to first response with VMP treatment was 1.4 months. Most first responses occurred early, but late de novo responses were seen, including 4% (9/238) that occurred beyond the first 24 weeks of therapy (ie, during cycles 5-9; Figure 3A). Quality of response improved with prolonged VMP treatment. Although most of patients’ best responses occurred within the first 24 weeks of therapy (cycles 1-4), a substantial proportion (16%, 37/238) occurred later (after 24 weeks), including 28% (29/102) of CRs (Figure 3B).
Among patients achieving CR with VMP by EBMT criteria, CR duration (time from CR to subsequent therapy) appeared similar in patients achieving CR early (cycles 1-4, within 24 weeks) versus late (cycles 5-9, after 24 weeks, Figure 4A). TNT (Figure 4B; median 41.2 vs 37.8 months), response duration (Figure 4C; median 44.7 vs 36.7 months), and OS (Figure 4D) appeared somewhat longer in the patients who achieved CR late versus early. Because they were measured from the time of randomization, TNT and OS appeared somewhat prolonged with late versus early CRs, with similar CR duration. Mean duration of therapy was 8.4 versus 7.6 cycles of study treatment in the patients achieving CR late versus early.
CR duration (time from CR to subsequent therapy) was analyzed among patients who received < 9 or all 9 protocol-specified cycles of bortezomib as part of VMP therapy, excluding those who stopped treatment early due to progression, and appeared similar in both groups (Figure 5A). This was also the case in an analysis restricted to those patients who achieved CR early (Figure 5B). Analyses of outcomes restricted to patients receiving all 9 cycles of bortezomib treatment demonstrated that CR duration (Figure 5C) appeared similar with late versus early achievement of CR, whereas TTP (data not shown), TNT (Figure 5D; median 46.2 vs 37.8 months), and response duration (data not shown) appeared somewhat longer in patients with late versus early CR, which was consistent with prior analyses (Figure 4). OS (data not shown) was similar between patients achieving late versus early CR in this subgroup receiving all 9 cycles of bortezomib.
Discussion

This analysis of the VISTA phase 3 trial demonstrated the differential prognostic significance of quality of response on outcomes in the nontransplant setting. The large VISTA population and high CR rate with VMP has provided a large sample size to reliably explore this prognostic significance. In a previous analysis, San Miguel et al\textsuperscript{16} showed that OS is superior with VMP vs MP alone; here, we show achievement of CR by EBMT response criteria was associated with significantly longer TTP, TNT, and TFI versus PR. However, there was no statistically significant difference in OS with CR versus PR. With the increasingly widespread use of highly active novel agent-based regimens in the relapsed setting, analyses of OS according to parameters or response in the front-line setting are becoming problematic due to the confounding effect of achieving a high-quality response to salvage therapy.\textsuperscript{21} In the VISTA population overall, 411 of 682 patients (60%) had received subsequent therapy by data cutoff of March 2009, including 52% and 69% of patients in the VMP and MP arms, respectively.\textsuperscript{17} Among patients on the VMP arm who had received subsequent therapy, 76% had received bortezomib-, thalidomide-, or lenalidomide-based therapy.\textsuperscript{17} Thus, in patients with inherently sensitive MM (that is, those who responded to VMP), additional responses to subsequent lines of therapy, particularly novel agent-based therapy, may have resulted in the similar OS observed between patients achieving CR or PR to VMP, despite the other significant time-to-event and clinical benefits of CR versus PR. In addition, improved outcomes were associated with > PR versus < PR, with patients achieving significantly longer TTP, TNT, TFI, and OS. These analyses, by both EBMT and IMWG criteria, suggest that quality of response is associated with improved long-term outcomes with VMP treatment. In addition, in a separate analysis of HRQoL data from VISTA, CR achievement and duration were associated with clinically relevant improvements in patient HRQoL.\textsuperscript{2}
These findings support the results of previous studies investigating the prognostic impact of CR versus PR on long-term outcomes in previously untreated MM patients, who were ineligible for, or did not undergo, HDT-SCT. In the ECOG E9486 trial, in which patients had active MM and received conventional chemotherapy, OS was 5.1 and 3.3 years in patients experiencing CR and PR, respectively (P < .0001). Similarly, in a retrospective analysis assessing clinical outcomes in 758 patients, Wang et al demonstrated longer OS after a 2-year landmark in patients who had achieved CR versus PR versus no response (12.4 years vs 3.7 years vs 2.1 years, respectively) at 2 years after receiving dexamethasone-based therapy.

Our analyses according to response by IMWG criteria showed improved outcomes with CR versus VGPR and with CR + VGPR versus PR. However, outcomes appeared similar with VGPR versus PR, indicating that there may not be a significant prognostic difference between VGPR and PR. This may be attributable to small numbers of patients achieving VGPRs (n = 41; 28 VMP, 13 MP) and these findings should be interpreted with caution. However, the superiority of CR over VGPR/PR is consistent with an analysis of outcomes and clinical benefit by quality of response in the bortezomib arm of the phase 3 APEX (Assessment of Proteasome Inhibition for Extending Remissions) trial, which demonstrated that CR was associated with significantly longer median TFI (24.1 vs 6.9/6.4 months) and time to alternative therapy (27.1 vs 13.6/14.0 months) versus VGPR/PR.

Two phase 2 studies of thalidomide-based regimens support our findings that CR + VGPR is superior to PR. Patients treated with pegylated liposomal doxorubicin, vincristine, dexamethasone, and thalidomide who achieved CR + VGPR demonstrated improved PFS (P = .01) and OS (P = .03) versus patients achieving PR/no response. Similarly, patients treated with thalidomide, dexamethasone, and pegylated liposomal doxorubicin who achieved CR + VGPR had a significantly higher 3-year TTP (78% vs 40%, P = .031), event-free survival (78% vs 37%, P = .021), and OS (84% vs 61%, P = .053) rates versus patients achieving < VGPR.

Our analyses have demonstrated that the clinical benefit of CR with VMP appears similar, regardless of whether CR is achieved early or late. A substantial proportion of patients achieved CR late in their treatment course (after 24 weeks); these so-called “late CRs” may include patients improving from an initial response of PR. Thus, our data support continuation of VMP therapy beyond a lesser response in patients tolerating treatment, with the aim of achieving a CR. A lower proportion of early versus late CRs were achieved in elderly patients, and patients with β2-microglobulin > 5.5 mg/L and low albumin (< 3.5 g/dL), suggesting that these patients take longer to achieve CR. In contrast, patients with ISS stage I had a greater proportion of early versus late CRs compared with ISS stage II and III patients, which suggests a more rapid complete response in patients with low tumor burden. In patients receiving all 9 cycles of bortezomib with VMP therapy, long-term outcomes appeared either similar or slightly improved with late versus early CR. This may be because of the inherent bias of the additional time taken to achieve a CR from a lesser response. These results support findings from the APEX trial, in which response duration was not affected by rapid achievement of first response in patients receiving bortezomib.

The use of techniques such as serum free light chain assay, and immunohistochemistry or immunofluorescence have led to a more stringent definition of CR (sCR by IMWG criteria) than immunofixation. Furthermore, multiparametric flow cytometry and polymerase chain reaction have allowed an even more stringent definition of CR through evaluation of minimal residual disease. Indeed, the greater depth of response defined by these more sensitive techniques has been shown to be prognostic for improved outcomes. Thus, a CR defined by these more sensitive techniques may be the ultimate treatment goal, rather than achievement of an immunofixation-negative CR.
In conclusion, these results demonstrate that CR is prognostic for improved long-term outcomes compared with lesser responses and so achievement of CR, by the most stringent definition possible, should be considered an important treatment goal. Thus, our data also support continuation of VMP therapy in patients tolerating treatment to achieve maximal response and maximal clinical benefit.

References


