Supplementary Online Content


eMethods. Inclusion/exclusion criteria, response criteria, MRD detection, FDG-PET/CT methodology, carfilzomib metabolism correlative methodology, statistical methods

eResults. MRD measurement results, subset analyses, pharmacokinetic analyses

eTable 1. Overall MFC, NGS, and PET responses in patients with NDMM or SMM

eFigure 1. Treatment schema

eFigure 2. Duration of response in patients with NDMM

eFigure 3. Pharmacokinetic analyses in patients with NDMM

eReferences

This supplementary material has been provided by the authors to give readers additional information about their work.
eMethods

Inclusion criteria for patients with newly diagnosed multiple myeloma (NDMM)

The protocol was open for newly diagnosed transplant/nontransplant-eligible patients with multiple myeloma (MM). Inclusion criteria consisted of clinical and histologically confirmed diagnosis of MM according to International Myeloma Working Group criteria\(^1\) with measurable disease (serum monoclonal protein (M-protein) ≥ 1 g/dL or urine M-protein >200 mg/24 hour or serum free light chain >10 mg/dL and skewed free light chain ratio), creatinine clearance ≥ 60 mL/min, age ≥ 18 years, Eastern Cooperative Oncology Group performance status 0–2, adequate hematologic parameters (absolute neutrophil count ≥ 1.0 K/uL, hemoglobin ≥ 8 g/dL, platelet count ≥ 75 K/uL), adequate hepatic function (bilirubin <1.5 \times upper limit of normal (ULN), aspartate transaminase (AST) and alanine transaminase (ALT) <3.0 ULN), ability to tolerate thombophrophylaxis, and participation in REMS\(^®\) (Celgene, Summit, NJ) programs.

Inclusion criteria for patients with smoldering MM (SMM)

Adult patients with confirmed SMM according to International Myeloma Working Group criteria\(^2\) were enrolled. Patients must have had measurable disease within the previous 4 weeks (defined as having either serum monoclonal protein of ≥ 3 g/dL, urinary Bence-Jones protein (BJP) level of >200 mg/24 hours, or a serum free light chain level >10 mg/dL, with an abnormal kappa/lambda ratio), and high-risk disease (per criteria from the Mayo Clinic\(^3\) or the Programa para el Estudio de la Terapéutica en Hemopatía Maligna [PETHEMA]\(^4\)). Other inclusion criteria included an Eastern Cooperative Oncology Group performance status ≤ 2, absolute neutrophil count ≥ 1000/μL, platelet count ≥ 75,000/μL, total bilirubin level ≤ 1.5 times the institutional ULN, AST–ALT ratio ≤ 2.5 times the institutional ULN, and creatinine clearance ≥ 60 mL/min.

Exclusion criteria

Exclusion criteria included prior treatment for MM (exception of steroids, bisphosphonates, radiotherapy for hypercalcemia, spinal cord compression, or aggressively progressing MM), plasma cell leukemia, pregnant or lactating females, uncontrolled hypertension or diabetes, active hepatitis infection, New York Heart Association class III or IV congestive heart failure or significant cardiovascular disease, refractory gastrointestinal disease, alternative uncontrolled illness, baseline grade ≥ 3 peripheral neuropathy, major surgery within 1 month, and contraindication to concomitant or supportive medications.

Response criteria

Response criteria were categorized according to International Myeloma Workshop Consensus Panel\(^1\) with the addition of near complete response (nCR) category\(^5\). Per criteria, 24-hour urine collections were used to confirm complete response (CR); for other categories, if needed, BJP quantification (mg/24 hour) was calculated following determination of the random urine BJP:creatinine ratio.\(^6,7\)
Multiparametric flow cytometry (MFC) methodology for minimal residual disease (MRD) detection

Bone marrow aspirate samples were processed within 12 hours of collection by prelysis of red cells, with $5 \times 10^6$ cells stained per tube, as previously. The eight-color immunophenotyping panel consisted of CD19 APC (clone SJ25C1), CD20APC-H7 (clone L27), CD27 PE (clone L128), CD38v450 (clone HB7), CD45 v500 (clone HI30), CD56 PE-Cy7 (clone NCAM16.2), CD81 FITC (clone JS-81), CD117 PE (clone 104D2), and CD138 PerCP-Cy5.5 (clone MI15) (BD Biosciences, San Jose, CA). In select cases in which the above panel was not sufficiently informative, intracellular kappa and lambda light chains (Polyclonal Rabbit Anti-Human, F(ab’)2, Dako) were examined in combination with CD19APC (clone SJ25C1), CD20APC-H7 (clone L27), CD38v450 (clone HB7), CD45v500 (clone HI30), CD56PE-Cy7 (clone NCAM16.2), and CD138 PerCP-Cy5.5 (clone MI15). Specimens were acquired using BD FACSCanto™ II (BD Biosciences, San Jose, CA) as previously described with an acquisition goal of $3 \times 10^6$ total events. Data were analyzed using FCS Express (De Novo Software, Los Angeles, CA), according to the European Myeloma Network Criteria. Specifically, doublet discrimination was performed, and plasma cells were identified by concurrent examination of CD138, CD38, CD45, and light scatter properties. Abnormal plasma cells were identified and distinguished from normal plasma cells, according to recommendations by the European Myeloma Network, which are consistent with the methods employed by PETHEMA, and as previously described. Specifically, expression patterns of CD19, CD20, CD27, CD38, CD45, CD56, CD81, CD117, and CD138 were used; in select cases, monoclonality was confirmed by intracellular light chain expression. A discrete population of $\geq 20$ abnormal plasma cells was necessary to define the presence of plasma cell disease. MRD testing in marrow aspirates was considered adequate when a population of abnormal plasma cells was detected. In the absence of demonstrable MRD, testing was considered acceptable when greater than $2 \times 10^6$ cells were acquired and the presence of marrow elements was immunophenotypically demonstrable (myeloblasts, precursor B cells, and/or normal plasma cells). Samples that were negative for MRD and in which marrow elements were not demonstrable or contained fewer than $2 \times 10^6$ cells were categorized as inadequate for MRD testing.

$^{18}$F-Fluorodeoxyglucose-positron emission tomography-computed tomography (FDG-PET/CT) methodology

$^{18}$F-FDG injection procedures were implemented according to standard institutional practice. Whole body (vertex to toes) static PET/CT imaging was performed beginning at 1-hour post injection. Focal lesions on $^{18}$F-FDG-PET/CT were identified using the following criteria: the presence of increased tracer (above background reference) within the bone, excluding articular processes irrespective of corresponding lesion identified on CT, maximum standardized uptake value (SUV) $>1.5$ for osteolytic lesions ranging from 0.5–1.0 cm, or maximum SUV $>2.5$ for lesions $>1.0$ cm. Two independent, blinded reviewers assessed focal lesions identified at baseline for post-therapeutic responses. Responses for $^{18}$F-FDG-PET/CT were adapted from Zamagni criteria and included the following categories: negative, defined as SUV $<1.5$ (lesions 0.5–1.0 cm) or SUV $<2.5$
(lesions >1.0 cm) or below background reference; decreased, defined as all baseline positive areas declining in mean percent change in SUV ≥25% but not meeting negative criteria; partial, defined as mixed response of negative, positive, and decreased FDG-avid lesions; and positive, defined as SUV ≥1.5 (lesions 0.5–1.0 cm) or SUV ≥2.5 (lesions >1.0 cm) or above background reference.

**Carfilzomib metabolism correlative methodology**

Carfilzomib clearance was compared with baseline laboratory values (albumin, alpha-globulin, β-2-microglobulin, white blood cell counts, basophils, absolute neutrophil count, lymphocytes, liver function tests, bilirubin, blood urea nitrogen, estimated creatinine clearance) that may have affected distribution or metabolism of carfilzomib in the blood, liver, or kidney. To study the rate of metabolism in plasma, apheresis samples were obtained from the National Institutes of Health blood bank, lyophilized albumin was added to plasma with and without a nonspecific protease inhibitor cocktail (1×). Carfilzomib concentrations were obtained at several time points over the course of 24 hours by mass spectrometry via previously published methods.

**Statistical methods for the NDMM study**

The study was a single arm phase 2 clinical trial intending to enroll a total of 45 evaluable patients, with the primary objective being to determine if the rate of grade ≥3 neurologic toxicity after two completed cycles was lower than 10%. The sample size selected was chosen to demonstrate that if the rate of grade ≥3 neurologic toxicity is lower than 10% or if it could be consistent with a rate that is greater than 10%, it would be considered excessive. The objective was to enroll a total of 45 evaluable patients in a single cohort, and score each patient for the development of grade ≥3 neurologic toxicity in the first two completed cycles of treatment. If the trial were to enroll 45 patients and if five or more had grade ≥3 neurologic toxicity, then the probability that the true rate of grade ≥3 neurologic toxicity of 5% would be 7.3%, and the probability that the true rate of grade ≥3 neurologic toxicity is 15% would be 82.5%. Thus, the appearance of five or more patients with grade ≥3 neurologic toxicity in 45 patients would have been adequate to provide evidence that the true rate of toxicity is consistent with 15% and considered excessive.

As an early stopping rule for toxicity for the first 20 patients, if four patients were found to have grade ≥3 neurologic toxicity in the first two completed cycles of treatment, no further patients were to have been enrolled as soon as this was determined, since the lower 80% one-sided confidence bound around four of 20 is 11.7%, which, being greater than 10%, would be considered demonstration of excessive toxicity.

Secondary endpoints included response rates, duration of response (DOR), and progression-free survival (PFS). PFS and DOR were estimated using the Kaplan-Meier method and reported along with 95% confidence intervals at appropriate time points. The significance of the difference between a pair of Kaplan-Meier curves was determined by an exact two-tailed log-rank test. Associations between two dichotomous parameters were determined with Fisher’s exact test while an exact Cochran-Armitage test for trend
was used to determine the association between an ordered categorical parameter and a dichotomous parameter.\textsuperscript{15} All $P$ values are two-tailed and are reported without any adjustment for multiple comparisons except when noted.

For pharmacokinetic or survival analysis, when comparisons were initially made based on data divided into quartiles and then subsequently divided at one of the quartiles, the resulting $P$ value was adjusted by multiplying the unadjusted $P$ value by the number of implicit tests used to decide to divide the data at that point.

Analyses were performed using commercially available software (SAS version 9.3 [SAS Institute, Cary NC] and StatXact 9 [Cytel Software Corporation, Cambridge MA]).

**Statistical methods for the SMM study**

The primary end point was the proportion of patients achieving a very good partial response (VGPR) or better after eight cycles of carfilzomib, lenalidomide, and dexamethasone (CRd). We selected a sample size of 12 patients for the study for the following reason: if the true probability of a VGPR was 20\% or 50\%, we calculated that there would be a 7.3\% or 80.6\% probability, respectively, that at least five patients would exhibit a VGPR. Thus, if at least five patients of 12 achieved a VGPR, there would be strong evidence that the true probability of a VGPR was $\geq$50\%.

Secondary endpoints included PFS, DOR, safety, and correlative assays assessing MRD.
eResults

Measuring MRD by MFC, NGS, and PET/CT

At baseline, all 57 (100%) patients demonstrated abnormal flow at baseline. Using NGS, ≥1 clonal rearrangement was identified in 43 of 46 (93%) of bone marrow CD138+ cell samples and in 46 of 57 (81%) of cell-free bone marrow aspirates with an overall clonal rearrangement detection rate of 49 of 57 (86%) at baseline. MRD assessment for NGS was done from cell-free supernatant BM aspirate. At MRD assessment (achievement of CR/sCR or end of eight cycles), one patient did not have a sample collected for evaluation of post-treatment residual disease (prior to completing eight cycles of CRd and had not achieved CR/sCR). MRD testing was clinically feasible in 55 of 56 (98%) patient samples by MFC; one patient sample was not processed due to equipment breakdown. Using the NGS method, eight patients did not have clonotype detected at baseline using either sample type and could not be compared with MRD sample time points (three CD138+ bone marrow cell samples/five cell-free bone marrow aspirates), and three patients lacked available samples to perform assay. Overall, NGS assay at MRD was technically successful in 45 of 56 (80%) patient samples.

PET/CT response intermediate category contains both decreased and partial PET/CT responders. Among 17 patients with NDMM completing 1 year of lenalidomide extension, one patient went from MFC-negative/PET-negative status to MFC-positive/PET-positive status (Patient 18) and another patient went from MFC-negative/PET-negative status to MFC-negative/PET-decreased status (Patient 4). Both remain clinically well, not meeting the definition of progressive disease despite FDG avidity increases in original identified lesions after 1 year of lenalidomide. Patient 2 had clinical progression on study when FDG-PET/CT detected a new FDG-avid lytic lesion after 1 year of lenalidomide extension (MFC-negative/PET-decreased status to MFC-negative/PET-positive status (note that the patient was NGS-positive after CRd therapy). All other patients retained their respective MFC status after 1 year of lenalidomide; 3 patients demonstrated FDG-PET/CT improvement (Patients 9 [partial to decreased], 15, and 16).

Subset analyses in patients with NDMM

In subset analyses of NDMM patients, with limited power to detect differences, best responses did not vary on the basis of age ≥65 years vs <65 years, unfavorable cytogenetics/fluorescence in situ hybridization (FISH), or presence of extramedullary disease at baseline (data not shown). PFS and DOR (≥partial response) were nondifferential by cytogenetics (PFS: P= .91; DOR: P=.25); PFS tended to be superior for patients aged ≥65 years compared with those <65 years (PFS: P=.03; DOR: P=.05). Based on seven cases of extramedullary disease at baseline, DOR was shorter in extramedullary-positive vs -negative cases (P=.04).

Discussion of subset analyses in patients with NDMM

© 2015 American Medical Association. All rights reserved.
Prior work suggests that certain genetic (ie, FISH) myeloma subtypes\textsuperscript{16,17} may adversely impact clinical outcomes but may be overcome with specific therapies.\textsuperscript{18} Although the current study was underpowered to evaluate this, based on small numbers and limited follow-up, outcomes of clinical response and PFS were not impacted by unfavorable FISH/cytogenetics in NDMM patients. In addition, population studies have shown improved survival outcomes in older patients,\textsuperscript{19} while specific melphalan-free regimens like continuous lenalidomide and dexamethasone have demonstrated improved PFS in transplant-ineligible patients compared with melphalan-based regimens.\textsuperscript{20} Our results demonstrated that patients aged $\geq 65$ years had improved DOR and PFS and similar responses to patients aged $<65$ years after CRd with good tolerability of therapy. While clinical trials may be subject to selection bias for older patients who tolerate therapy well, such a consideration only highlights that fitness may be a better outcome predictor than age.

**Pharmacokinetic analyses in patients with NDMM**

We observed that three progressors with available pharmacokinetic data had carfilzomib clearance values (975.91, 2146.7, and 5013.6 L/h) greater than two-fold over median (528.4 L/h) clearance. No direct statistical significant association between carfilzomib clearance and clinical outcome was found; however, to gain an understanding of these outcomes, baseline laboratory results that could be associated with drug metabolism/elimination were assessed in relation to carfilzomib clearance. Baseline albumin was related to carfilzomib clearance (exact Jonckheere-Terpstra trend-test, $P=.016$); the lower three quartiles of serum albumin had consistent clearance values (median=442.6, 95% confidence interval [CI]: 404.4–699.4 L/h) while those in the fourth quartile (albumin $>3.9$ g/dL) had abnormally high and variable clearance (median=1576 L/h, 95% CI: 964.6–3387 L/h), (Wilcoxon rank sum test, unadjusted $P=.014$, adjusted $P=.042$) (eFigure 3A). Estimated 18-month PFS for patients with baseline albumin $\leq 3.9$ g/dL (quartiles 1–3) vs $>3.9$ g/dL (fourth quartile) was 100% vs 67.5% (95 CI: 36.3%–88.3%) (exact two-tailed log-rank, unadjusted $P=.0034$; adjusted $P=.010$) (eFigure 3B). Additional analyses demonstrated trends that showed that higher baseline albumin concentration was associated with poorer clinical responses (<nCR/CR/stringent CR, MRD positivity by MFC and next-generation sequencing) (eFigure 3C). To confirm these findings, albumin was added to human apheresis plasma. It was found to increase carfilzomib metabolism ex vivo; conversely, nonspecific protease inhibitor decreased carfilzomib metabolism and reduced the metabolic rate in plasma that contained albumin (eFigure 3D).

**Discussion of pharmacokinetic analyses in patients with NDMM**

As part of the study design, pharmacokinetic data were collected on patients, and carfilzomib clearance was noted to be elevated among a few clinical progressors. Prior work established that carfilzomib metabolism is primarily mediated through extrahepatic peptidase cleavage. Along these lines, we found albumin $>3.9$ g/dL are associated with higher carfilzomib clearances and counterintuitively suboptimal clinical outcomes (shorter PFS, MRD-positivity). In further support, our ex vivo experiments confirmed that albumin addition to plasma potentiated carfilzomib metabolism and the potentiation.
was reversible using nonspecific protease inhibitor. We found no statistical association between carfilzomib clearance and clinical outcomes; however, the analysis was limited by small numbers and short follow-up. Based on available knowledge, our interpretation is that the therapeutic impact of carfilzomib is dictated by multiple elements, including host and tumor biology and drug metabolism.


**Table 1.** Overall MFC, NGS, and PET Responses in Patients With NDMM or SMM

<table>
<thead>
<tr>
<th>MRD by MFC among nCR/CR/sCR(^a)</th>
<th>NDMM</th>
<th>SMM</th>
<th>NDMM + SMM</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative</td>
<td>28/28 (100)</td>
<td>11/12 (92)</td>
<td>39/40 (98)</td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>0/28 (0)</td>
<td>1/12 (8)</td>
<td>1/40 (3)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MRD by MFC among ≤VGPR(^a)</th>
<th>NDMM</th>
<th>SMM</th>
<th>NDMM + SMM</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative</td>
<td>5/15 (33)</td>
<td>—</td>
<td>5/15 (33)</td>
<td>&lt;.001(^b)</td>
</tr>
<tr>
<td>Positive</td>
<td>10/15 (67)</td>
<td>—</td>
<td>10/15 (67)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MRD by NGS among nCR/CR/sCR(^c)</th>
<th>NDMM</th>
<th>SMM</th>
<th>NDMM + SMM</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative</td>
<td>14/21 (67)</td>
<td>9/12 (75)</td>
<td>23/33 (70)</td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>7/21 (33)</td>
<td>3/12 (25)</td>
<td>10/33 (30)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MRD by NGS among ≤VGPR(^c)</th>
<th>NDMM</th>
<th>SMM</th>
<th>NDMM + SMM</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative</td>
<td>0/12 (0)</td>
<td>—</td>
<td>0/12 (0)</td>
<td>&lt;.001(^b)</td>
</tr>
<tr>
<td>Positive</td>
<td>12/12 (100)</td>
<td>—</td>
<td>12/12 (100)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PET among nCR/CR/sCR(^d)</th>
<th>NDMM</th>
<th>SMM</th>
<th>NDMM + SMM</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative</td>
<td>11/27 (41)</td>
<td>—</td>
<td>—</td>
<td>.11(^e)</td>
</tr>
<tr>
<td>Decrease</td>
<td>7/27 (26)</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Partial</td>
<td>9/27 (33)</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>0/27 (0)</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PET among ≤VGPR(^d)</th>
<th>NDMM</th>
<th>SMM</th>
<th>NDMM + SMM</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative</td>
<td>4/16 (25)</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Decrease</td>
<td>4/16 (25)</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Partial</td>
<td>5/16 (31)</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>3/16 (19)</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: CR, complete response; nCR, near complete response; MFC, multiparametric flow cytometry; MRD, minimal residual disease; NDMM, newly diagnosed multiple myeloma; NGS, next-generation sequencing; PET, positron emission tomography; sCR, stringent complete response; SMM, smoldering multiple myeloma; VGPR, very good partial response.

\(^a\)Two patients with NDMM were not evaluable for MFC or PET response (data not collected).

\(^b\)By Fisher’s exact test using NDMM + SMM.

\(^c\)Twelve patients with NDMM were not evaluable for NGS (data not collected). No SMM patients obtained a ≤VGPR rate.

\(^d\)Overall designated as best responses after a median potential follow-up of 17.3 months (patients with NDMM) or 16 months (patients with SMM).

\(^e\)By the exact Cochran-Armitage test.
eFigure 1. Treatment schema

Each cycle is 28 days. For patients <75 years, stem cells were harvested after at least four cycles of CRd and continued with combination therapy after harvesting.

*Carfilzomib dose of 20 mg/m² on C1D1 and D2, and 36 mg/m² thereafter. Carfilzomib infusions were 30 minutes. bDexamethasone dose of 20 mg in cycles 1–4 and 10 mg in cycles 5–8. Abbreviations: C, cycle; CRd, carfilzomib, lenalidomide, and dexamethasone; D, day.
eFigure 2: Duration of response in patients with NDMM.

(A) Duration of PR. (B) Duration of CR/sCR. Abbreviations: CR, complete response; NDMM, newly diagnosed multiple myeloma; PR, partial response; sCR, stringent complete response.
eFigure 3: Pharmacokinetic analyses in patients with NDMM

(A) Carfilzomib clearance by albumin quartile (B) PFS by albumin quartiles (lower three quartile ≤3.9 g/dL vs fourth quartile albumin >3.9 g/dL) (C) Albumin quartiles and clinical endpoints. (D) Addition of albumin affects carfilzomib metabolism in plasma ex vivo. Abbreviations: ALB, albumin; CI, confidence interval; CR, complete response; nCR, near complete response; MFC, multiparametric flow cytometry; MRD, minimal residual disease; NGS, next-generation sequencing; PFS, progression-free survival; PI, proteasome inhibitor; sCR, stringent complete response; VGPR=very good partial response.
eReferences


© 2015 American Medical Association. All rights reserved.


