

Published in final edited form as:

Cancer. 2010 February 1; 116(3): 640–646. doi:10.1002/cncr.24767.

Residual Serum Monoclonal Protein Predicts Progression-Free Survival in Patients With Previously Untreated Multiple Myeloma

Eric W. Schaefer, MS¹, Shaji Kumar, MD², Angela Dispenzieri, MD², Jacob B. Allred, MS¹, Morie A. Gertz, MD², Martha Q. Lacy, MD², S. Vincent Rajkumar, MD², and Sumithra J. Mandrekar, PhD¹

¹Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, Minnesota

²Division of Hematology, Mayo Clinic, Rochester, Minnesota

Abstract

BACKGROUND—Currently used treatment response criteria in multiple myeloma (MM) are based in part on serum monoclonal protein (M-protein) measurements. A drawback of these criteria is that response is determined solely by the best level of M-protein reduction, without considering the serial trend. The authors hypothesized that metrics incorporating the serial trend of M-protein would be better predictors of progression-free survival (PFS).

METHODS—Fifty-five patients with measurable disease at baseline (M-protein ≥ 1 g/dL) who received ≥ 4 cycles of treatment from 2 clinical trials in previously untreated MM were included. Three metrics based on the percentage of M-protein remaining relative to baseline (residual M-protein) were considered: metrics based on the number of times residual M-protein fell within prespecified thresholds, metrics based on area under the residual M-protein curve, and metrics based on the average residual M-protein reduction between Cycles 1 and 4. The predictive value of these metrics was assessed in Cox models using landmark analysis.

RESULTS—The average residual M-protein reduction was found to be significantly predictive of PFS ($P = .02$; hazard ratio, 0.37), in which a patient with a 10% lower average residual M-protein reduction from Cycle 1 to 4 was estimated to be at least 2.7 \times more likely to develop disease progression or die early. None of the other metrics was predictive of PFS. The concordance index for the average residual M-protein reduction was 0.63, compared with 0.56 for best response.

CONCLUSIONS—The average residual M-protein reduction metric is promising and needs further validation. This exploratory analysis is the first step in the search for treatment-based trend metrics predictive of outcomes in MM.

Keywords

multiple myeloma; prediction; progression-free survival; response; serum monoclonal protein

Traditionally, efficacy of antineoplastic therapies has been measured in terms of changes in physical dimensions for the majority of solid tumors and for hematologic malignancies, assessments have been based on change in tumor cell counts or tumor-secreted markers. In the context of myeloma, availability of an easily measurable tumor marker in the form of serum or urine monoclonal protein has allowed easy estimates of the tumor burden in the majority of

© 2009 American Cancer Society.

Corresponding author: Sumithra J. Mandrekar, PhD, PhD, Department of Health Sciences Research, Mayo Clinic, 200 First Street SW, Rochester, MN 55905; Fax: (507) 266-2477; mandrekar.sumithra@mayo.edu.
Eric W. Schaefer's current address: Division of Biostatistics, Penn State College of Medicine, Hershey, Pennsylvania.

patients.^{1,2} Currently used treatment response criteria have been developed based on serum or urine monoclonal protein measurements using protein gel electrophoresis and immunofixation studies. The advent of the serum free light chain assay in recent times has allowed for assessment based on changes in the serum free light chain for patients lacking an intact immunoglobulin in the serum.³ Long-term implications of the degree of response obtained by any specific therapy have been evaluated in multiple clinical trials using different treatment regimens. In the setting of myeloma, attainment of a very good partial response (PR) or complete response (CR) has always translated into improved progression-free survival (PFS), but the effect on overall survival has been less consistent.⁴⁻⁷ This is likely a reflection of the underlying disease biology as well as the specific therapies used.

One of the drawbacks of the traditional response criteria is that they are solely determined by the baseline and the best level of reduction in the tumor estimate obtained with a specific therapy, with little emphasis given to the trend in reduction of the tumor marker. Some of the published data suggest that a rapid response to therapy predicts rapid progression and a poor overall outcome, particularly in the context of changes in free light chain levels.⁸ Unlike high-dose therapy and autologous stem cell transplantation, in which most of the tumor reductions occur relatively rapidly after the transplant, continuous cyclical treatment regimens are associated with an incremental decrease over time. This is particularly true for the newer immunomodulatory agents, in which gradual restoration or enhancement of the antitumor immunity might lead to a steady and consistent decrease in tumor burden over the duration of therapy.⁹⁻¹⁰ The currently used European Group for Blood and Marrow Transplant Response Criteria and the International Uniform Response Criteria are based on best percentage reduction in monoclonal protein (M-protein). These need to be examined to determine whether incorporation of measures that assess the serial trend of response will allow further improvements and better prediction of longer-term outcome.^{1,2} We undertook this investigation in the context of 2 clinical trials that used newer treatment agents, namely thalidomide and lenalidomide, in combination with dexamethasone.

MATERIALS AND METHODS

Patients

This research study was approved by the Mayo Clinic Institutional Review Board. Patient data were combined from 2 clinical trials (988013 and MC038D), in which previously untreated multiple myeloma patients were enrolled. In 988013, the regimen was thalidomide and dexamethasone; in MC038D, the regimen was lenalidomide and dexamethasone. Inclusion/exclusion criteria, treatment schedules, and other details have been previously published.¹¹⁻¹² Both trials demonstrated antitumor activity. Patients with measurable disease at baseline (serum M-protein ≥ 1.0 g/dL) who received at least 4 cycles of therapy were included in this analysis. We used this cutoff because the response at 4 months was considered the primary endpoint in these 2 trials, and patients were allowed to go to stem cell transplantation after 4 cycles if they so desired.

Metrics

We were interested in identifying alternative metrics for predicting PFS outcomes for patients with measurable serum monoclonal protein using electrophoresis that would be an improvement over currently used response criteria or add to the existing criteria. The European Group for Blood and Marrow Transplant Response Criteria¹ was originally used to define response and progression in both studies and was retained for the current analysis. Response categories were CR, very good PR, PR, minor response, and stable disease (SD). For this analysis, we categorized patients into 2 groups on the basis of response in the first 4 cycles of treatment: patients with best response of CR or very good PR versus patients with best response

of PR, minor response, or SD. This categorization was based on the consistent relation seen in previous studies between CR/very good PR and PFS, and the small numbers of patients in the different response categories (CR = 1; very good PR = 7; PR = 38; minor response = 2; SD = 7). All metrics constructed in the analysis were compared with this dichotomous best response variable.

The metrics constructed were based on serial M-protein measurements obtained at baseline and after each cycle. For easier interpretability, the percentage of M-protein remaining relative to the baseline level was computed at each cycle. We denote this measure as residual M-protein. For example, a patient with a baseline M-protein of 2.5 g/dL and a post-Cycle 1 M-protein of 1.2 g/dL would have an residual M-protein of 48% after Cycle 1. The metric residual M-protein was analyzed at each cycle as a simple metric by itself. We also defined 3 trend metrics based on residual M-protein, discussed below. A representation of these metrics using data from a hypothetical patient is given in Table 1.

The first metric was a categorical metric using the number of times residual M-protein fell within specified thresholds from Cycle 1 to 4 (see Metric 1 in Table 1). For the categorical trend metric, if 2 of the 4 nonbaseline residual M-protein values were below a specified threshold, then the patient was classified as “low”; if 2 of the 4 nonbaseline residual M-protein values were above a different specified threshold, then the patient was classified as “high”; otherwise, the patient was classified as “intermediate.” Two separate classifications were used: 1 in which the lowest threshold was 33% and the largest threshold 66%, and 1 in which the lowest was 20% and the largest 60%. These metrics are referred to as 33/33/33 and 20/40/40, respectively. The approach offered a simplistic way to measure the residual M-protein trend.

The second metric was a continuous metric for the area under the residual M-protein curve from baseline to a particular cycle (denoted by AUC_p, for area under the residual M-protein curve). The AUC_p metric was intended to measure the cumulative effect of the residual M-protein trend (see Metric 2 in Table 1). This metric was calculated by constructing line segments between consecutive residual M-protein values, and obtaining the area under the resulting curve. Patients with a more gradual reduction in residual M-protein over time had a larger AUC_p than patients with a sharper reduction. AUC_p was computed for Cycles 2, 3, and 4, but not for Cycle 1 because it was equivalent to residual M-protein at Cycle 1.

The third metric, the average residual M-protein reduction from Cycle 1 to 4, was constructed as follows: for the set of consecutive cycles, the residual M-protein at a particular cycle was subtracted from the residual M-protein in the previous cycle, and the resulting differences averaged across the set (see Metric 3 in Table 1). For example, the hypothetical patient shown in Table 1 has residual M-protein of 48%, 28%, 40%, and 16% from Cycle 1 to 4, respectively. The corresponding residual M-protein reductions are 20%, -12%, and 24%, resulting in an average residual M-protein reduction of 11%. The metric can be equivalently constructed by taking the difference in residual M-protein from Cycle 1 to 4 and dividing by 3, which is the number of treatment cycles. Positive values denote a decrease in residual M-protein, and negative values denote an increase. The residual M-protein reduction from Cycle 0 to 1 was equivalent to residual M-protein at Cycle 1, and the residual M-protein reduction from Cycle 0 to 4 was equivalent to residual M-protein at Cycle 4, and hence not considered as individual metrics.

Statistical Analysis

Cox proportional hazards models¹³ were used to assess the predictive ability of the metrics for PFS using a landmark analysis. Specifically, the Cycle 4 evaluation date was used as the baseline. PFS was defined as the time from the date of the Cycle 4 evaluation until the date of progression or death from any cause. Patients alive and progression-free at the date of last

follow-up were censored. To adjust for the finding that each study had separate but similar inclusion/exclusion criteria, all models included a stratum effect for study. *P* values <.05 were considered evidence of significantly predictive metrics, though no adjustments were made to control for the inflated false-positive rate, as this was an exploratory analysis. The concordance index¹⁴ was used to estimate the predictive ability of each metric, with a value of 0.50 reflecting no predictive ability, and a value of 1 reflecting perfect predictive ability. A bootstrap technique¹⁵ was used to identify whether metrics selected on the basis of *P* values showed evidence of being false positives. Specifically, a stratified bootstrap approach was used such that for each bootstrap sample, the same numbers of patients were selected from each study as in the original sample. The *P* value and concordance index were recorded from each of 1000 bootstrap samples. The percentage of bootstrap samples in which a metric had a *P* value <.05 was reported. In addition, 95% confidence intervals (95% CIs) were constructed using the Bca method¹⁶ for the difference in concordance index between the metrics and best response, to determine whether the metric had superior predictive ability in comparison to best response.

Two multivariate models were analyzed in the same fashion as above. The first model included the residual M-protein reduction from Cycle 0 to 1 and the average residual M-protein reduction from Cycle 1 to 4, and aimed to decompose the residual M-protein trend into its initial and sharp reduction and the trend after that initial reduction. The second model included the average residual M-protein reduction from Cycle 1 to 4 and residual M-protein at Cycle 4, and aimed to measure both the trend over 4 cycles and the tumor burden at Cycle 4.

RESULTS

Patient Characteristics

Across both studies, a total of 68 previously untreated multiple myeloma patients had an M-protein by serum protein electrophoresis ≥ 1 g/dL at baseline. Thirteen (19%) patients did not receive a full 4 cycles of treatment and were excluded; the final sample size is 55 patients, with 31 patients from 988013 and 24 patients from MC038D.

Patient characteristics are presented in Table 2. Thirty-seven (67%) events (disease progression or death from any cause) were observed, 27 (87%) in 988013 and 10 (42%) in MC038D. The median follow-up time from the end of Cycle 4 for the 18 censored patients was 1.97 years (range, 43 days to 3.07 years). Patients in MC038D had a longer median PFS from the end of Cycle 4 (2.75 years to 1.83 years), although the difference was not statistically significant (log-rank *P* = .14).

Characteristics of Residual M-Protein Trends

A sample of representative residual M-protein trends is shown in Figure 1. The trends can be loosely characterized into 4 groups, from most common to least common: 1) an initial sharp decrease followed by a gradual decrease through the remaining cycles; 2) a sharp initial decrease followed by either a plateau or a slight increase; 3) a complete or nearly complete disappearance of M-protein from baseline to Cycle 1 followed by slight variation in subsequent cycles; and 4) a gradual reduction through all cycles. These trends are a potentially important source of variation, and the metrics aim to measure features of these trends.

Assessment of Metrics

Descriptive statistics for each metric are shown in Table 3. Results from the Cox proportional hazards models are presented in Table 4. Estimates and hazard ratios (HRs) for continuous metrics are reported for meaningful units, roughly $\frac{1}{2}$ of the range of the metric. A forest plot of the estimated HRs is shown in Figure 2. The only metric that was significantly predictive of PFS was the average residual M-protein reduction from Cycle 1 to 4 (*P* = .02; HR = 0.37),

in which a patient with a 10% lower average residual M-protein reduction was estimated to be 2.70× more likely to have an earlier event (ie, progression or death). However, only 66% of the bootstrapped samples had a *P* value <.05 for this metric. On the basis of the concordance index, the average residual M-protein reduction from Cycle 1 to 4 and the residual M-protein at Cycle 4 had superior predictive ability (concordance index = 0.63) compared with all other metrics, including best response (concordance index = 0.56).

The 95% Bca CIs for the difference in concordance index as compared with the best response model were -0.05 to 0.18 for the average residual M-protein reduction from Cycle 1 to 4, and -0.02 to 0.18 for the residual M-protein at Cycle 4. This indicates that there is not enough evidence to definitively conclude that these metrics have significantly better predictive ability than best response alone, possibly because of the small sample size.

DISCUSSION

This limited pooled analysis attempted to identify metrics predictive of PFS in previously untreated multiple myeloma patients. Most metrics, including best response, showed little ability to predict PFS. The average residual M-protein reduction from Cycle 1 to 4 was the only metric with a significant *P* value, and also had the largest concordance index. Residual M-protein at Cycle 4 also had the same concordance index, but the *P* value was non-significant. Residual M-protein at Cycle 1 was not a significant predictor of PFS, a surprising result because most patients had a sharp initial reduction, and the reduction is the most striking feature of the residual M-protein trend. Both multivariate models based on the average residual M-protein reduction from Cycle 1 to 4 showed no advantage over the univariate counterpart. For these reasons, the average residual M-protein reduction from Cycle 1 to 4 seems to be most promising.

The results of the current analysis are not conclusive. The sample size is limited, and any proposed metric would need to be strongly predictive of PFS for statistical significance. Only 66% of the bootstrap replicates had a significant *P* value for the average residual M-protein reduction. Although the concordance index reflected moderate discriminate ability, the predictive ability of this metric in comparison to best response was not statistically superior. A larger sample size is necessary for any definitive conclusions.

The average residual M-protein reduction from Cycle 1 to 4 offers several advantages in comparison to best response. First, this metric uses data from multiple cycles accounting for the serial trend. Second, this metric can be easily extended to include more cycles and translated to other time units with little effort. Third, although the interpretation of the metric is not as straightforward as best response, the average reduction in residual M-protein is qualitatively easy to understand. One drawback is that this metric does not consider all of the data, specifically the reduction from baseline to cycle 1. However, our analysis showed that this initial reduction had no predictive value.

There are limitations to this analysis. The exploratory procedure of creating and assessing the metrics may have led to spurious conclusions, as there was no reasonable way to control for the false-positive rate in this small study cohort of 55 patients. Data were collected for regimens containing only the immunomodulatory drugs thalidomide and lenalidomide. Other chemotherapy regimens, such as ones containing proteasome inhibitors, may not exhibit the same characteristics of the residual M-protein trend and in turn eliminate the predictive ability of these metrics.

A validated metric based on the first few cycles of treatment could have potentially important ramifications for clinical trial design. The median PFS for previously untreated multiple myeloma patients undergoing an autologous stem cell transplantation is approximately 18

months, with many of the newer regimens when used as primary therapy giving comparable results.^{17–19} Many phase 2 trials are designed with a binary endpoint of PFS rate at a particular time point such as 1 year. This has clear drawbacks: the time point selected is largely arbitrary, although the expected number of events is taken into consideration when choosing the time point; the design relies on historical control rates, which may be difficult to obtain for certain populations or treatment regimens; and 1 year of follow-up to obtain the primary endpoint results may still result in a lengthy phase 2 trial. A validated metric predictive of PFS, such as the average residual M-protein reduction proposed in this manuscript, may offer an acceptable alternative endpoint for phase 2 trials. Karrison et al²⁰ describe a randomized design for continuous endpoints using treatment and control arms. Such a design would require 46 total patients (23 per arm) to test the hypothesis that the treatment arm would have a 5% higher average residual M-protein reduction through the first 4 cycles compared with the control arm, with 90% power at the 5% significance level, assuming a common standard deviation of 5% (as seen in this analysis). The advantages of using a randomized phase 2 design with average residual M-protein reduction as the endpoint are: 1) the sample size for this phase 2 trial is similar to using the design with historical control rates; 2) the design uses a control arm to assess the activity of the treatment regimen, alleviating the need to generate historical control rates; 3) the primary endpoint is based on measurements from the first few cycles and thus can be assessed quickly; and 4) most importantly, the results of the trial would be expected to predict the more clinically important endpoint of PFS. Admittedly, any metric used in these designs would need to be more fully validated.

In summary, the results of the current study indicate that there exists potential for using serial M-protein measurements from the first few cycles of treatment to predict PFS in previously untreated multiple myeloma patients with measurable M-protein. We recognize that prospective analyses using 1) more patients, 2) a larger variety of drug regimens, 3) other assessment methods (ie, urine protein electrophoresis, free light chain levels), and 4) impact on overall survival are needed to validate these initial findings. Nevertheless, to our knowledge, this is a first step in the search for treatment-based trend metrics predictive of PFS in multiple myeloma.

Acknowledgments

CONFLICT OF INTEREST DISCLOSURES

Supported by Mayo Clinic Comprehensive Cancer Center grant CA-15,083.

REFERENCES

1. Blade J, Samson D, Reece D, et al. Criteria for evaluating disease response and progression in patients with multiple myeloma treated by high-dose therapy and haemopoietic stem cell transplantation. Myeloma Subcommittee of the EBMT European Group for Blood and Marrow Transplant. *Br J Haematol* 1998;102:1115–1123. [PubMed: 9753033]
2. Durie BGM, Harousseau J-L, Miguel JS, et al. International uniform response criteria for multiple myeloma. *Leukemia* 2006;20:1467–1473. [PubMed: 16855634]
3. Katzmann JA, Clark RJ, Abraham RS, et al. Serum reference intervals and diagnostic ranges for free kappa and free lambda immunoglobulin light chains: relative sensitivity for detection of monoclonal light chains. *Clin Chem* 2002;48:1437–1444. [PubMed: 12194920]
4. Kyle RA, Leong T, Li S, et al. Complete response in multiple myeloma: clinical trial E9486, an Eastern Cooperative Oncology Group study not involving stem cell transplantation. *Cancer* 2006;106:1958–1966. [PubMed: 16565956]
5. Rajkumar SV, Fonseca R, Dispenzieri A, et al. Effect of complete response on outcome following autologous stem cell transplantation for myeloma. *Bone Marrow Transplant* 2000;26:979–983. [PubMed: 11100277]

6. Lonial S, Gertz MA. Eliminating the complete response penalty from myeloma response assessment. *Blood* 2008;111:3297–3298. [PubMed: 18332233]
7. Alexanian R, Weber D, Giralt S, et al. Impact of complete remission with intensive therapy in patients with responsive multiple myeloma. *Bone Marrow Transplant* 2001;27:1037–1043. [PubMed: 11438818]
8. van Rhee F, Bolejack V, Hollmig K, et al. High serum-free light chain levels and their rapid reduction in response to therapy define an aggressive multiple myeloma subtype with poor prognosis. *Blood* 2007;110:827–832. [PubMed: 17416735]
9. Lacy M, Gertz M, Dispenzieri A, et al. Lenalidomide plus dexamethasone (Rev/Dex) in newly diagnosed myeloma: response to therapy, time to progression, and survival. *Blood* 2006;108:798.
10. Richardson PG, Sonneveld P, Schuster MW, et al. Bortezomib or high-dose dexamethasone for relapsed multiple myeloma. *N Engl J Med* 2005;352:2487–2498. comment 2546–2548. [PubMed: 15958804]
11. Rajkumar SV, Hayman S, Gertz MA, et al. Combination therapy with thalidomide plus dexamethasone for newly diagnosed myeloma. *J Clin Oncol* 2002;20:4319–4323. [PubMed: 12409330]
12. Rajkumar SV, Hayman SR, Lacy MQ, et al. Combination therapy with lenalidomide plus dexamethasone (Rev/Dex) for newly diagnosed myeloma. *Blood* 2005;106:4050–4053. [PubMed: 16118317]
13. Cox DR. Regression models and life-tables (with discussion). *J R Stat Soc Series B Stat Methodol* 1972;34:187–220.
14. Harrell, FE, Jr. *Regression Modeling Strategies With Applications to Linear Models, Logistic Regression, and Survival Analysis*. New York, NY: Springer-Verlag; 2001.
15. Efron, B. *The Jackknife, the Bootstrap and Other Resampling Plans*. Philadelphia, PA: Society for Industrial and Applied Mathematics; 1982.
16. Carpenter J, Bithell J. Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. *Stat Med* 2000;19:1141–1164. [PubMed: 10797513]
17. Attal M, Harousseau JL. Standard therapy versus autologous transplantation in multiple myeloma. *Hematol Oncol Clin North Am* 1997;11:133–146. [PubMed: 9081208]
18. Child JA, Morgan GJ, Davies FE, et al. High-dose chemotherapy with hematopoietic stem-cell rescue for multiple myeloma. *N Engl J Med* 2003;348:1875–1883. [PubMed: 12736280]
19. Femand JP, Katsahian S, Divine M, et al. High-dose therapy and autologous blood stem-cell transplantation compared with conventional treatment in myeloma patients aged 55 to 65 years: long-term results of a randomized control trial from the Group Myelome-Autogreffe. *J Clin Oncol* 2005;23:9227–9233. [PubMed: 16275936]
20. Karrison T, Maitland M, Stadler W, et al. Design of phase II cancer trials using a continuous endpoint of change in tumor size: application to a study of sorafenib and erlotinib in non-small-cell lung cancer. *J Natl Cancer Inst* 2007;99:1455–1461. [PubMed: 17895472]

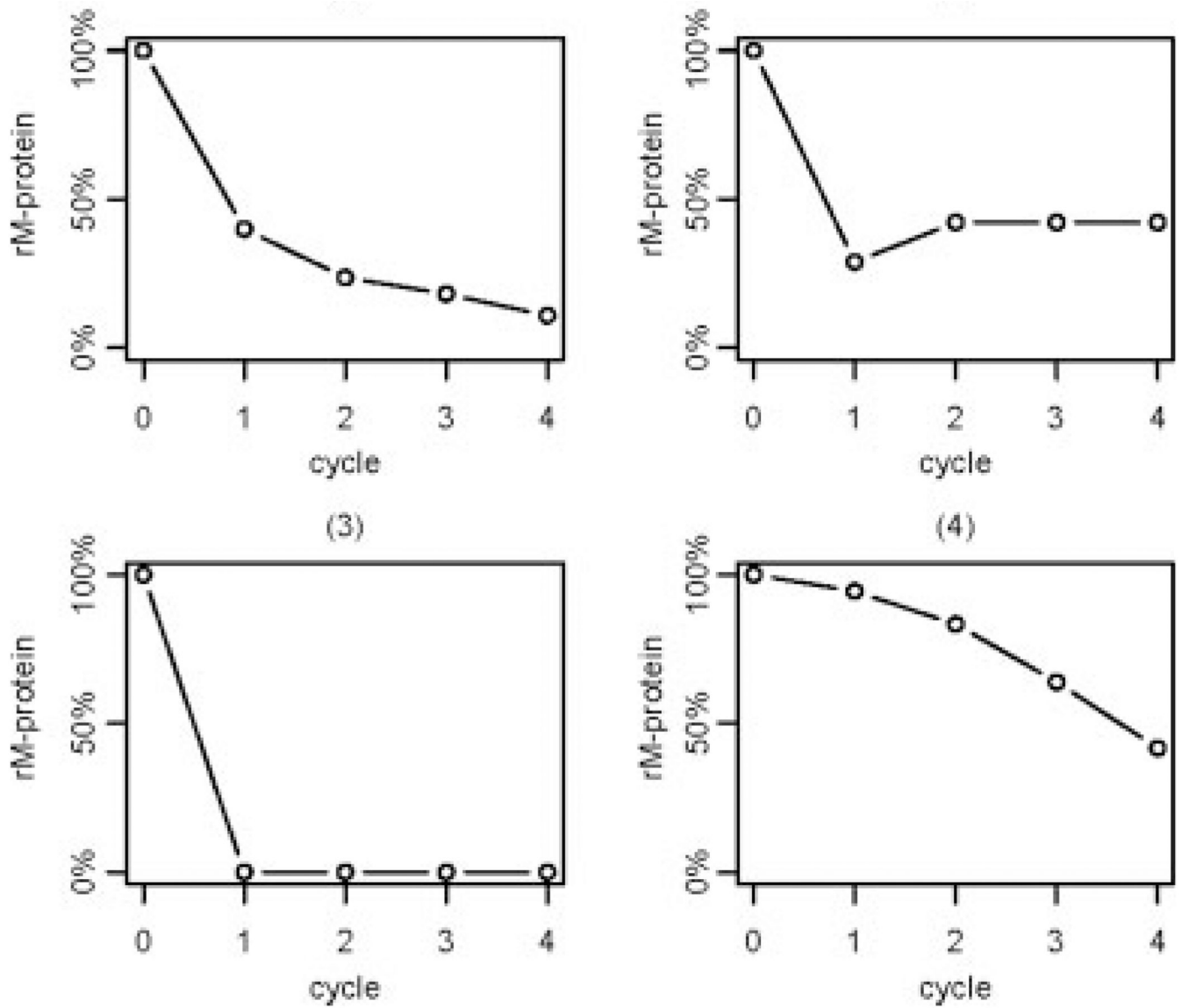


Figure 1.

A representative sample of residual monoclonal protein (rM-protein) trends through the first 4 cycles of treatment is shown. rM-protein indicates the percentage of serum monoclonal protein remaining relative to the baseline level.

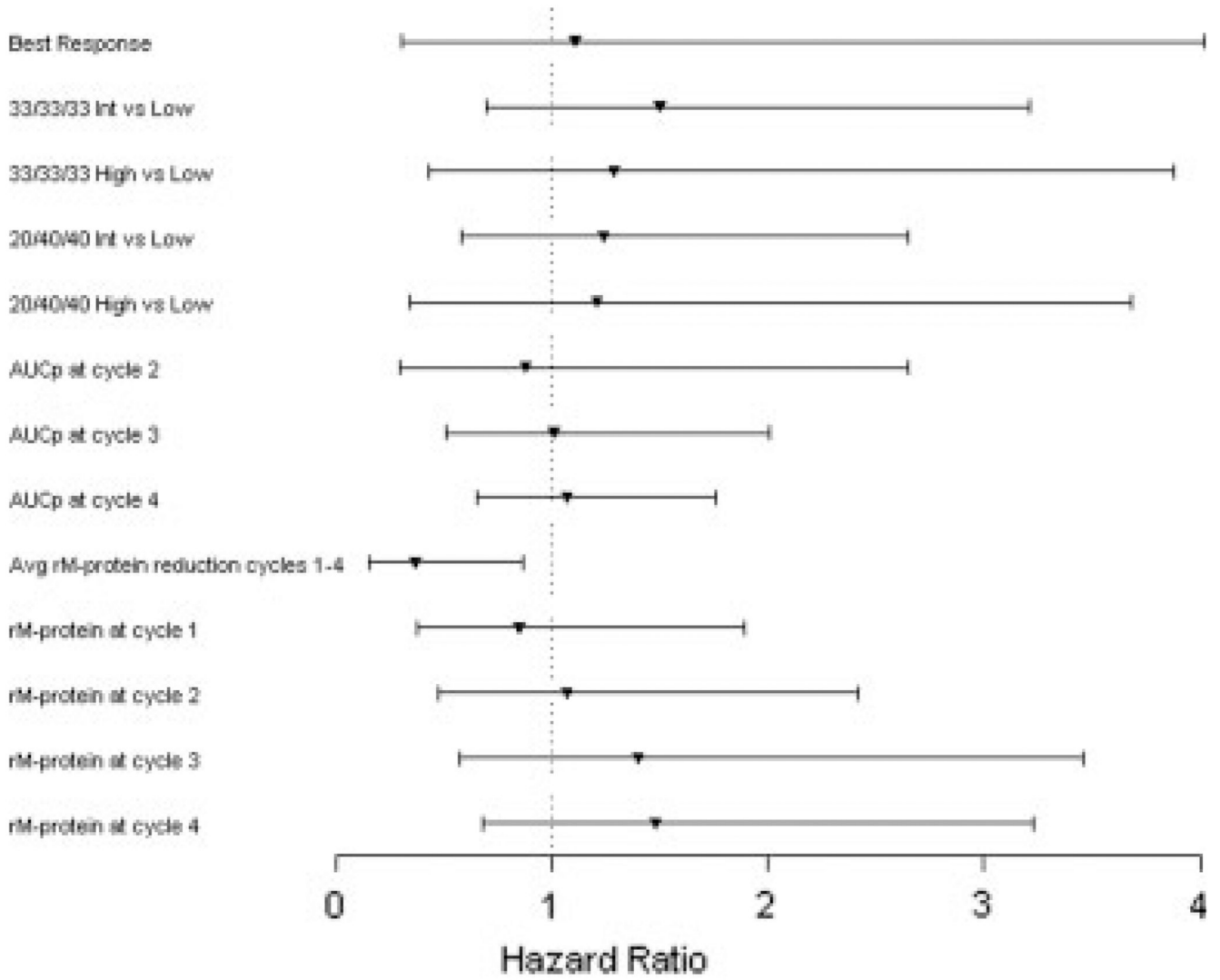


Figure 2. A forest plot is shown of estimated hazard ratios and 95% Wald confidence intervals for each metric residual monoclonal protein (rM-protein), which is the percentage of serum monoclonal protein remaining relative to the baseline level. Int indicates intermediate; AUCp, area under the rM-protein curve from baseline to a particular cycle; Avg, average.

Table 1

Calculation of Metrics for a Hypothetical Patient

M-protein	g/dL
Baseline	2.5
Cycle 1	1.2
Cycle 2	0.7
Cycle 3	1.0
Cycle 4	0.4
rM-protein	%
Cycle 1	48
Cycle 2	28
Cycle 3	40
Cycle 4	16
Metric 1	
33/33/33	Low
20/40/40	Intermediate
Metric 2	%
AUCp	
Cycle 2	112
Cycle 3	146
Cycle 4	174
Metric 3	%
Avg rM-protein reduction Cycles 1–4	11

M-protein indicates monoclonal protein; rM-protein, residual M-protein (the percentage of serum monoclonal protein remaining relative to the baseline level); AUCp, area under the rM-protein curve from baseline to a particular cycle; Avg, average.

Table 2

Patient Characteristics

Characteristic	988013 [N531]	MC038D [N524]
Median age, y (range)	61 (38–75)	64 (40–78)
Men	19 (61%)	15 (63%)
ECOG Performance status		
0	17 (55%)	10 (42%)
1	11 (36%)	12 (50%)
2	2 (6%)	2 (8%)
3	1 (3%)	0 (0%)

ECOG indicates Eastern Cooperative Oncology Group.

Table 3

Descriptive Statistics for Each Metric

Continuous Metrics	Median (SD)	25th–75th Percentile	Minimum-Maximum
AUCp at Cycle 2	109% (34%)	87%–133%	50%–193%
AUCp at Cycle 3	142% (55%)	103%–179%	50%–274%
AUCp at Cycle 4	174% (75%)	113%–215%	50%–361%
Avg rM-protein reduction in Cycles 1–4	4% (5%)	2%–7%	–6% to 19%
rM-protein at Cycle 1	42% (22%)	28%–56%	0%–103%
rM-protein at Cycle 2	37% (25%)	20%–54%	0%–83%
rM-protein at Cycle 3	26% (22%)	11%–41%	0%–83%
rM-protein at Cycle 4	28% (22%)	8%–43%	0%–93%
Categoric Metrics	CR/VGPR	PR/MR/Stable Disease	
Best Response	8 (15%)	47 (85%)	
	Low	Intermediate	High
33/33/33	32 (58%)	16 (29%)	7 (13%)
20/40/40	21 (38%)	25 (45%)	9 (16%)

SD indicates standard deviation; AUCp, area under the residual monoclonal protein (rM-protein) curve from baseline to a particular cycle; Avg, average; rM-protein, the percentage of serum monoclonal protein remaining relative to the baseline level; CR, complete response; VGPR, very good partial response; PR, partial response; MR, minor response.

Table 4

Cox Proportional Hazards Model Results

Models	Unit	Estimate (SE)	Hazards Ratio (95% Wald CI)	P (Boot %)	c-Index
Univariate models					
Best response	—	0.11 (0.66)	1.11 (0.31–4.02)	.87 (3%)	0.56
33/33/33	—			.58 (15%)	0.57
Intermediate vs low		0.41 (0.39)	1.50 (0.70–3.21)		
High vs low		0.25 (0.56)	1.29 (0.43, 3.88)		
20/40/40	—			.86 (8%)	0.53
Intermediate vs low		0.21 (0.39)	1.24 (0.58–2.65)		
High vs low		0.11 (0.61)	1.21 (0.34–3.68)		
AUCp at Cycle 2	100%	-0.12 (0.56)	0.88 (0.30–2.65)	.83 (4%)	0.52
AUCp at Cycle 3	100%	0.01 (0.35)	1.01 (0.51–2.01)	.97 (4%)	0.55
AUCp at Cycle 4	100%	0.07 (0.25)	1.07 (0.65–1.76)	.78 (5%)	0.58
Avg rM-protein reduction in Cycles 1–4	10%	-0.99 (0.43)	0.37 (0.16–0.87)	.02 (66%)	0.63
rM-protein at Cycle 1	50%	-0.17 (0.41)	0.85 (0.38–1.89)	.68 (5%)	0.50
rM-protein at Cycle 2	50%	0.07 (0.41)	1.07 (0.48–2.42)	.86 (4%)	0.55
rM-protein at Cycle 3	50%	0.34 (0.46)	1.40 (0.57–3.46)	.46 (12%)	0.59
rM-protein at Cycle 4	50%	0.39 (0.40)	1.48 (0.68–3.23)	.32 (19%)	0.63
Multivariate models					
rM-protein reduction in Cycles 0–1	10%	-0.03 (0.08)	0.97 (0.83–1.13)	.66 (8%)	0.62
Avg rM-protein reduction in Cycles 1–4	10%	-1.05 (0.35)	0.35 (0.14–0.85)	.02 (68%)	
Avg rM-protein reduction in Cycles 1–4	10%	-0.95 (0.44)	0.39 (0.16–0.92)	.03 (58%)	
rM-protein at Cycle 4	50%	0.17 (0.40)	1.19 (0.54–2.60)	.66 (9%)	0.62

SE indicates standard error; 95% CI, 95% confidence interval; c-Index, concordance index; AUCp, area under the residual monoclonal (rM-protein) curve from baseline to a particular cycle; Avg, average; rM-protein, the percentage of serum monoclonal protein remaining relative to the baseline level.