

Appraisal of immunoglobulin free light chain as a marker of response

Angela Dispenzieri,^{1,2} Lijun Zhang,³ Jerry A. Katzmann,² Melissa Snyder,² Emily Blood,⁴ Roberta DeGoey,¹ Kimberly Henderson,¹ Robert A. Kyle,^{1,2} Martin M. Oken,⁵ Arthur R. Bradwell,⁶ and Philip R. Greipp^{1,3}

Departments of ¹Hematology and ²Laboratory Medicine, Mayo Clinic, Rochester, MN; ³Department of Biostatistics, Eastern Cooperative Oncology Group (ECOG), Boston, MA; ⁴Dartmouth Medical School, Lebanon, NH; ⁵University of Minnesota, Minneapolis; and ⁶Department of Immunology, Birmingham University Medical School, Birmingham, United Kingdom

The immunoglobulin free light chain (FLC) assay is an invaluable tool for following patients with oligosecretory plasma cell dyscrasia. Baseline values have also been shown to be prognostic in all plasma cell disorders tested. A looming question, however, is the role it should play in following myeloma patients with disease that is measurable using serum and urine electrophoresis. We used the data and stored samples from a mature Eastern Cooperative Oncology Group clinical trial

(E9486) to assess serum levels of FLC at baseline and after 2 months of alkylator-based therapy. For serial determinations, the absolute level of involved serum FLC or the difference of the involved and uninvolved FLC is preferred over the ratio of involved to uninvolved FLC. FLC response after 2 months of therapy was superior to early M-protein measurement to predict overall response. The ideal cut-point for FLC change appears to be between 40% and 50% reduction. The

correlation between serial measurements of serum FLC and urine M-protein is inadequate to abolish the serial 24-hour urine protein. Although baseline values of FLC are prognostic in newly diagnosed myeloma patients, serial measurements do not appear to have added value in patients who have M-proteins measurable by electrophoresis. (Blood. 2008;111:4908-4915)

© 2008 by The American Society of Hematology

Introduction

The free light chain (FLC) assay is a nephelometric measurement of kappa and lambda light chains that circulate as light chain monomers or dimers and that are not bound to immunoglobulin heavy chain. The quantitation of the kappa and lambda FLC and the calculation of the FLC κ/λ ratio are sensitive and specific for detection of excess monoclonal FLC.¹ Studies from populations of patients with nonsecretory multiple myeloma (MM),² light chain amyloidosis,³⁻⁶ light chain deposition disease,^{1,4} monoclonal gammopathy of uncertain significance,⁷⁻⁹ plasmacytoma,¹⁰ smoldering MM,¹¹ light chain MM,¹² and newly diagnosed MM^{13,14} have documented the sensitivity and prognostic utility of these assays and established their use as a complement to immunofixation.

In addition to its diagnostic use, the assay is being used for monitoring disease course in light chain amyloidosis, light chain deposition disease, nonsecretory MM, and light chain MM in which there may be a band detected on immunofixation that cannot be quantitated by protein electrophoresis.^{3,5,6} Empiric definitions for “free light chain response” have been proposed for patients with amyloidosis in the Consensus Opinion from the 10th International Symposium on Amyloid and Amyloidosis¹⁵ and for patients with MM in the International Uniform Response Criteria,¹⁶ but have not been validated. Several questions remain regarding the use of FLC in patients with MM. First, what is the best definition of FLC response? Second, does early FLC decline predict best overall response, early paraprotein response, progression-free survival, or overall survival in patients with Bence Jones myeloma and/or those patients with intact immunoglobulin? Third, can serum FLC measurement replace the 24-hour urine protein electrophoresis as a measure of light chain response?

Methods

Patients

From February 1988 to May 1992, 653 patients from 36 Eastern Cooperative Oncology Group (ECOG) institutions were enrolled in treatment trial E9486, and their outcomes have been previously reported.¹⁷ A total of 407 patients had matched sets of pretreatment and 2-month postserum samples available to run immunoglobulin FLC assays; 8 of these had nonsecretory disease and were excluded from this study, leaving 399 patients for the present study. Informed consent was obtained in accordance with the Declaration of Helsinki.

Eligibility and treatment details have been previously described.¹⁷ In brief, to enroll in the treatment trial, patients were required to have previously untreated myeloma and measurable disease defined as either a serum M-protein greater than or equal to 10 g/L or urine monoclonal light chain excretion exceeding 200 mg in 24 hours or serially measurable soft tissue plasmacytomas or bone marrow plasmacytosis greater than or equal to 20%. Patients entering the study were stratified according to age, creatinine, M-protein type, and clinical stage, and randomized to 1 of the 3 study treatment regimens. The first regimen was VBMCP (vincristine, carmustine [BCNU], melphalan, cyclophosphamide, and prednisone), administered in 5-week cycles. The second regimen was VBMCP plus recombinant alpha-2 interferon. Patients younger than 70 years were also randomized to a third regimen, VBMCP and high-dose cyclophosphamide. Treatment was continued for a maximum of 2 years or until disease progression if the latter occurred first.

Data collection

The baseline clinical data and overall response data are housed in the ECOG database and have been reported previously.¹⁷ The individual patient M-protein data after 2 cycles were abstracted onto data forms from original

Submitted February 10, 2008; accepted March 17, 2008. Prepublished online as *Blood* First Edition paper, March 25, 2008; DOI 10.1182/blood-2008-02-138602.

The publication costs of this article were defrayed in part by page charge

payment. Therefore, and solely to indicate this fact, this article is hereby marked “advertisement” in accordance with 18 USC section 1734.

© 2008 by The American Society of Hematology

patient paper forms (by M.S.), including only patients whose 2-month posttreatment research serum sample and urine sample were obtained before institution of a subsequent cycle. Two cycles of VBMCP with or without alpha-2 interferon were 70 days, and 2 cycles of VBMCP alternating with high-dose cyclophosphamide were 56 days.

Monoclonal protein measurements

The measurements for serum M-protein and urine M-protein were prospectively performed using serum and urine electrophoresis at each of the test centers enrolling patients, and these results were abstracted prospectively. The FLC assay (FREELITE, The Binding Site, Birmingham, United Kingdom) was retrospectively and centrally performed on a Dade Behring (Deerfield, IL) BNII automated nephelometer using 0.5 mL of stored thawed serum. This assay consists of 2 separate measurements: one to quantitate kappa FLC and the other to quantitate lambda FLC. In addition to reporting the kappa and lambda FLC, the assay report also contains the FLC κ/λ ratio (diagnostic range, 0.26-1.65). Patients with ratios more than 1.65 contain excess kappa FLC and are presumed to be producing clonal kappa FLC. Patients with ratios less than 0.26 contain excess lambda FLC and are presumed to be producing clonal lambda FLC.

Definitions

The involved FLC (iFLC) is defined as the actual value of serum immunoglobulin kappa FLCs in patients with monoclonal kappa plasma cells or of serum immunoglobulin lambda FLCs in patients with clonal lambda plasma cells. Two definitions for "measurable FLC" were tested: (1) an iFLC of greater than 0.1 g/L (10 mg/dL)^{15,16} and (2) an abnormal kappa-to-lambda ratio with a difference between involved and uninvolved FLC (dFLC) of at least 0.5 g/L (5 mg/dL). Seventy-one percent of patients satisfied the former definition and 80% the latter. Results were not significantly different, except where mentioned, and for the purpose of reporting the latter definition was used. Criteria for measurable disease by electrophoresis in the serum are an M-protein greater than or equal to 0.01 g/L and in the urine are greater than or equal to 200 mg/24 hours.¹⁷

According to original clinical trial design, an objective response based on electrophoretic measurements required a 50% decrease in serum M-protein or, in patients lacking a serum M-protein, a 90% decrease in 24-hour urine M-protein. These criteria were the standard ECOG response criteria of period, rather than the European Group for Blood and Marrow Transplant¹⁸ or the international uniform response criteria of the present day.¹⁶ For FLC response, we used aspects of the definition published by Gertz et al¹⁵ and Durie et al¹⁶. An FLC response was defined as follows: A, a 50% decrease in the dFLC¹⁶; or B, a 50% decrease in the level of iFLC (B1)¹⁵ AND a 50% decrease in the ratio of involved to uninvolved FLC (rFLC; B2). The major difference between our application of this definition and its original intent is that FLC response is usually applied to those patients without measurable disease by electrophoretic methods; all of our patients had measurable disease by serum and/or urine electrophoresis.

Statistical methodology

For the purpose of analyses, the 399 patients were divided into subgroups. A priori they were parsed by how their disease was measurable by standard criteria¹⁷: measurable disease in their serum and urine (serum and urine patients, $n = 241$), in their serum only (serum-only patients, $n = 103$), or in their urine only (urine-only patients, $n = 55$). All analyses were done for the total group as well as for each of these subgroups. For most analyses, patients were also split into groups based on whether they had normal ($n = 341$) or abnormal ($n = 58$) renal function defined as baseline serum creatinine greater than 2 mg/dL.

We evaluated each FLC response (conditions A, B1, and B2) for association with ECOG best overall response, 2-month paraprotein response, overall survival (OS), and progression-free survival (PFS). For each criterion, a Fisher exact test of proportions was performed to test for the association between FLC response after 2 cycles of treatment and standard electrophoretic response measurements. Two electrophoretic measurement endpoints were used. The first was paraprotein response after 2 months

(contemporaneous to FLC measurement), and the second was the best overall response (using standard ECOG response criteria for objective response). Survival comparisons (OS and PFS) were performed using the 2-sided log rank test between FLC responders and nonresponders. The effect of change in FLC level between baseline and 2 months on OS and PFS was also modeled using Cox proportional hazards models, corrected for other prognostic factors (serum albumin, plasma cell labeling index, and beta-2 microglobulin), and stratified by treatment assigned.

To compare FLC assessment and monoclonal protein assessment for predicting ECOG overall objective response status, the performance of the 2 methods was compared based on the sensitivity and specificity for predicting ECOG overall objective response status. The risk value, or expected loss, which is a function of sensitivity and specificity and associated with each test, was used for comparison. If the risk with one test was less than for the other, then there was evidence that the one with lower risk was preferred. In the analysis, only the patients with both paraprotein and FLC information at 2 months are included, and equal weight was given for false positive and false negative.¹⁹

To determine whether serum immunoglobulin FLC measurement could replace 24-hour urine M-protein (ie, urine light chain determination), the relationship of their behavior after 2 months of therapy compared with baseline was evaluated using the Spearman correlation. Percentage change of 24-hour M-spike and of 24-hour urine total protein was compared with percentage change of serum immunoglobulin FLC. This analysis was performed separately for patients with involvement in serum and urine patients, urine-only patients, and both. The analyses have been performed on the dFLC level, the iFLC level, and the rFLC (conditions A, B1, and B2).

Results

Patient characteristics

The characteristics of the 399 patients included in this study are shown in Table 1. The median age was 63.1 years, and 65% of patients were male. Sixty percent of patients had both measurable serum and urine M-proteins, whereas 26% had only a serum M-protein and 14% had only a urine M-protein. There were no differences among the patient groups with the exception of a lower International Staging System (ISS) risk score in the serum-only patients ($P < .001$) and higher iFLC levels in the urine-only group ($P < .001$). Ninety-six percent of patients had an abnormal kappa-to-lambda FLC ratio.

In toto, 72% of patients eventually achieved an objective ECOG partial response as defined by M-protein reductions, and only one patient was considered unevaluable (Table 2). The median time to response in responding patients was 3.4 months. The median PFS was 2.6 years (95% confidence interval [CI], 2.3-2.7), and the median overall survival was 3.7 years (95% CI, 3.4-4.0). Only one patient was lost to follow-up, and 95% have died during the monitoring period.

Does baseline FLC predict for overall or progression-free survival?

Baseline elevations in iFLC predicted for OS and PFS. An outcome-oriented method using log likelihood value was applied to find the best break point for predicting OS was done for the dFLC, and the best breakpoint was the value that coincided with the upper limit of the lowest tertile values. Patients were divided into tertiles; and regardless of the baseline variable used (ie, iFLC, dFLC, or rFLC), the patients with lowest tertile of FLC had the best outcomes compared with the 2 higher tertiles. The outcomes for the 2 higher tertiles were nearly identical. The OS and PFS curves for baseline dFLC are shown in Figure 1. The respective median OSs

Table 1. Baseline characteristics for all patients

Variable	Total, n = 399	Serum and urine, n = 241	Serum only, n = 103	Urine only, n = 55
Percentage of total	100%	60%	26%	14%
Male, %	64.7	68.0	63.1	52.7
Age, y	63.1 (23.8-83.5)	63.7 (34.7-83.5)	62.7 (23.8-76.2)	59.3 (32.5-82.1)
Serum creatinine, mg/dL	1.2 (0.5-4.9)	1.3 (0.5-4.9)	1.1 (0.7-2.2)	1.3 (0.6-4.7)
Serum heavy chain, %				
IgG	58.6	68.1	68.0	0
IgA	24.6	27	32.0	0
IgD	0.2	0.8	0	0
none	13.8	0	0	100
missing	2.5	4.2	0	0
Serum M-spike				
n	337	235	102	0
Median (range), g/dL	4.0 (1.0-10.0)	4.1 (1.0-10.0)	3.9 (1.4-9.4)	NA
Light chain clonality κ/λ , %	66.4/33.6	65.6/34.4	68.9/31.1	65.6/34.5
Serum FLC				
n	399	241	103	55
Involved, mg/dL	37.2 (0.25-3370)	47.9 (0.75-3290)	7.21 (0.25-642)	286 (7.49-3370)
Uninvolved, mg/dL	0.66 (0.01-13.1)	0.76 (0.01-12.1)	0.65 (0.02-13.1)	0.44 (0.04-1.77)
Difference of involved – uninvolved, mg/dL	36 (0.03-3369)	47.1 (0.4-3285)	6.04 (0.03-641)	284 (7.26-3368)
Ratio of involved/uninvolved	69.0 (1.03-28406)	82.6 (1.04-28 406)	13.6 (1.03-10 000)	613 (17.4-17 400)
Involved more than 10 mg/dL, %	71	78	41	98
Difference of involved minus uninvolved more than 5 mg/dL, %	80	86	53	100
ISS, %				
I	47.1	40.7	64.1	43.6
II	25.3	27.0	23.8	21.8
III	27.6	32.4	11.8	34.6
Serum β_2 -microglobulin, mg/L	3.6 (0.6-64)	4.1 (0.6-64)	2.7 (0.8-13.4)	4.2 (0.6-16.9)
Urine M-spike				
n	227	173	0	54
Median (range), g/24 hr	0.94 (0.01-22.7)	0.59 (0.01-13.7)	NA	3.13 (0.05-22.7)
BM plasma cells				
n	389	234	102	53
Median (range), %	40 (2-99)	44 (2-99)	35 (4-99)	43 (6-99)

Data are medians (range), except where specified. There were no differences among the groups with the exception of a lower ISS risk in the serum-only patients and higher involved FLC levels in the urine-only group.

NA indicates not applicable.

were 49.4, 42, and 42 months, and the median PFSs were 34.9, 28.7, and 29.5 months.

What happens to FLC after 2 months of therapy?

Not all patients were evaluable for FLC response because baseline levels were barely elevated. Eighty percent had an abnormal ratio with a dFLC of greater than 0.05 g/L, and their changes after 2 months of therapy are shown in Table 2. The median percentage change in the absolute value of iFLC, rFLC, and dFLC was a reduction of 58% for all definitions. The greatest FLC reductions were seen in the urine-only group, and the largest discrepancy among results for the 3 definitions was in the serum-only group; the percentage reduction using rFLC in the serum-only group was different from the percentage reduction using iFLC or dFLC.

Do early changes in FLC predict for response?

After 2 months of therapy, patients who had a serum immunoglobulin FLC response were more likely to demonstrate significant reduction in their serum and urine M-proteins as measured by electrophoresis both at the same time point (Table 3). However, only 23% (32 of 139) had achieved paraprotein response at this early time point. In contrast, 62% (86 of 139) had achieved an FLC response. The majority of early FLC responders converted into eventual

overall responders 85% (155 of 183). Stated slightly differently, 69% of the overall responders had an early FLC response.

The same analyses were repeated splitting patients into those with normal and abnormal renal function. The effect that FLC response after 2 cycles of chemotherapy had on eventual overall response was less strong in the 49 patients with baseline creatinine levels greater than 2 mg/dL (data not shown). In this group, 81% of the overall responders had an early FLC response, and 83% of the early FLC responders had an eventual overall response.

By comparing the risk value, which is a function of sensitivity and specificity for overall ECOG response prediction, the FLC response performed better than the paraprotein response at 2 months for the overall population (risk value = 0.30 and 0.48, respectively, $P = .001$) and in those with measurable baseline FLC (Table 4).

What is the minimal level of iFLC to be deemed “evaluable” for response?

We considered 2 thresholds for baseline “evaluable” FLC: iFLC greater than 0.1 g/L and dFLC greater than 0.05 g/L in the setting of an abnormal FLC ratio. Using the first definition, 71% of the patients were considered “evaluable” compared with 80% when using the second definition. The sensitivity, specificity, negative

Table 2. Hematologic responses and outcomes

	All patients	Serum and urine	Serum only	Urine only
Best overall response, %				
n	399	241	103	55
Objective response	71.7	70.5	72.8	74.5
No change	15.8	17.0	14.6	12.7
Progression	12.3	12.4	11.7	12.7
Unevaluable	0.3	0	1	0
Months to objective response (among responders), median (range)	3.4 (0.5-133.9)	3.4 (0.5-27.8)	2.6 (1-133.9)	4.1 (1.1-21.8)
Died, n (%)	95.0	97.5	90.3	92.7
FLC evaluable subset*				
n	318	208	55	55
Died, %	97	98	95	93
ECOG objective response, %	70	70	67	74
Months to objective response (among responders), median (range)	3.5 (0.5-27.8)	3.5 (0.5-27.8)	3.5 (1.1-25.2)	4.1 (1.1-21.8)
Responses after 2 months*				
% change of iFLC	-58.1 (-99.8-9953)	-56.7 (-99.8-2233)	-55.9 (-96-122)	-64.1 (-98.9-9953)
% change of rFLC	-57.7 (-99.9-9328)	-58.5 (-99.9-1068)	-34.0 (-99.4-1082)	-66.0 (-99.5-9328)
% change of dFLC	-58.4 (-319-10 263)	-57.9 (-319-2410)	-58.2 (-105-127)	-64.4 (-99.8-10 263)
FLC responder, definition A, %	58	56	62	62
FLC responder, definition B, %	43	41	42	49
Paraprotein responder, % (n/N)	23 (32/139)	22 (14/63)	25 (11/44)	22 (7/32)

*Including only those patients with measurable baseline FLC, that is, dFLC of greater than or equal to 5 mg/dL. Except where stated, results are expressed as median (range).

†Percent change equals = 100 × (2-month – baseline)/baseline.

predictive value, and positive predictive value of FLC response predicting for eventual overall ECOG response were comparable whether the baseline definition of “evaluable” was dFLC greater than 0.05 g/L or the iFLC greater than 0.1 g/L. The characteristics of these 36 patients who were FLC “evaluable” only by the second definition were as follows: one had a creatinine greater than or equal to 176.8 μmol/L; among the 35 remaining patients with creatinine less than 176.8 μmol/L, 20 were serum and urine patients, 14 were serum-only patients, and 1 was a urine-only patient. In this subset, 17 were considered FLC responders and 18 were nonresponders.

What is the best definition for free light chain response?

The individual components of FLC response and the overall definition were applied. The FLC response rates after 2 months of therapy using definitions A (A, a 50% decrease in dFLC) and B (B, a 50% decrease iFLC AND a 50% decrease in the rFLC) were 58% and 43%, respectively (Table 2). In 85% of patients, there was agreement between definitions A and B, with definition A responders capturing all definition B responders.

By dichotomizing each FLC measurement at 50% reduction, we explored which would be most predictive of ultimate overall

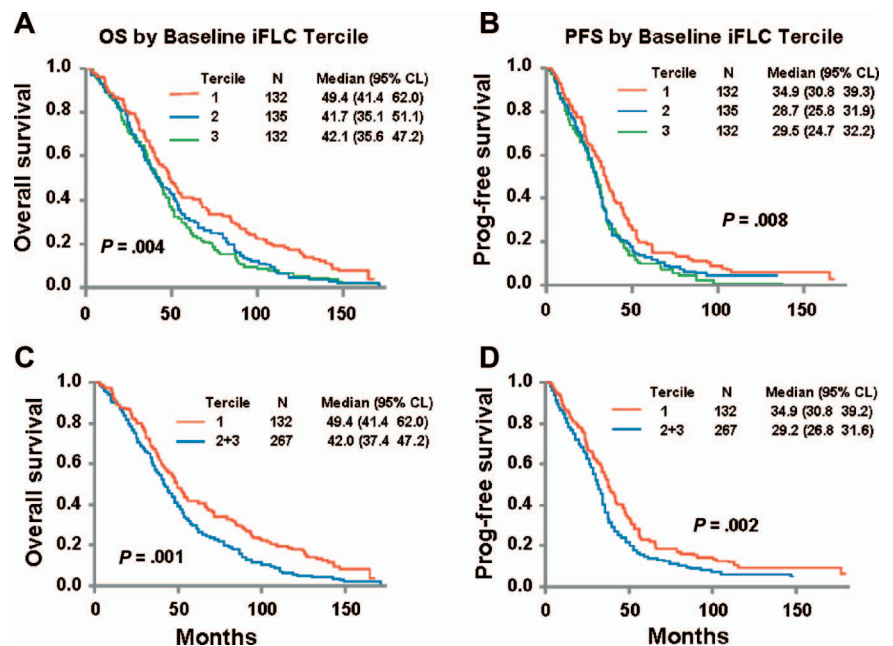


Figure 1. Overall survival is dependent on baseline FLC values. Tertile 1 is 0.03 to 11.77 mg/dL; tertile 2 is 11.77 to 85.56 mg/dL; tertile 3 is 85.6 to 3368.5 mg/dL (n = 399). (A) Overall survival: baseline involved minus uninvolved FLC. (B) Progression-free survival: baseline involved minus uninvolved FLC. (C) Overall survival: baseline involved minus uninvolved FLC (tertile 1 vs tertile 2 + tertile 3). (D) Progression-free survival: baseline involved minus uninvolved FLC (tertile 1 vs tertile 2 + tertile 3). Data shown are results for differences between baseline involved and uninvolved FLC; however, results are similar regardless of whether absolute value or ratio of involved to uninvolved FLC is used.

Table 3. Relationship of 2-month FLC response to standard M-protein response criteria

	FLC responders, no. (%)	Non-FLC responders, no. (%)	<i>P</i>
Association between 2-month FLC response* and 2-month paraprotein† response in patients with abnormal baseline FLC‡			
All patients			
Paraprotein responders	29 (91)	3 (9)	<.001
Nonparaprotein responders	57 (53)	50 (47)	
Serum and urine			
Paraprotein responders	13 (93)	1 (7)	.005
Nonparaprotein responders	25 (51)	24 (49)	
Serum only			
Paraprotein responders	10 (91)	1 (9)	.031
Nonparaprotein responders	17 (52)	16 (48)	
Urine only			
Paraprotein responders	6 (86)	1 (14)	.37
Nonparaprotein responders	15 (60)	10 (40)	
Association between 2-month FLC response and ECOG overall objective response in patients with abnormal baseline FLC‡			
All patients			
Overall objective responders	155 (69)	69 (31)	<.001
Nonresponders	28 (30)	65 (70)	
Serum and urine			
Overall objective responders	99 (68)	47 (32)	<.001
Nonresponders	17 (27)	45 (73)	
Serum only			
Overall objective responders	26 (70)	11 (30)	.070
Nonresponders	7 (41)	10 (59)	
Urine only			
Overall objective responders	30 (73)	11 (27)	.061
Nonresponders	4 (29)	10 (71)	

*An FLC response is defined here as a 50% decrease in the difference between involved and uninvolved FLC levels.

†Limited data for serum and urine electrophoresis at 2-month time point reduce sample size to 139 patients.

‡Abnormal FLC is defined as abnormal ratio and baseline iFLC greater than or equal to 5 mg/dL. Results were similar if all patients included or if abnormal FLC was defined as patients with baseline involved FLC value of greater than or equal to 10 mg/dL (data not shown).

ECOG response in patients with intact measurable serum immunoglobulin: iFLC, rFLC, or dFLC. There were no significant differences among the 3 measurements, but the risk value of dFLC was the lowest. Moreover, when we compared the FLC response definition A to definition B, definition A performed better than definition B (risk = 0.31 vs 0.36, *P* = .009).

To further test the definition of FLC response, a receiver operation curve was constructed to discern the best cutoff for FLC reduction after 2 months of therapy to predict for best overall response (Figure 2). A dFLC reduction of 40% to 50% had the best performance. A 50% reduction had a sensitivity of 69% and specificity of 70%.

Table 4. Comparison of FLC and paraprotein methods predicting ECOG overall objective response status*

FLC response	No. of patients	2-month FLC response			2-month paraprotein response			<i>P</i> †
		Sensitivity %	Specificity %	Risk	Sensitivity %	Specificity %	Risk	
Overall population	181	69	73	0.30	34	98	0.48	<.001
Patients with iFLC more than 10 mg/dL	120	83	67	0.24	36	100	0.43	.001
Patients with abnormal rFLC and dFLC 5 mg/dL or more	139	77	69	0.26	34	100	0.45	<.001

*Among patients with both paraprotein and FLC data after 2 months of therapy.

†Risk comparison.

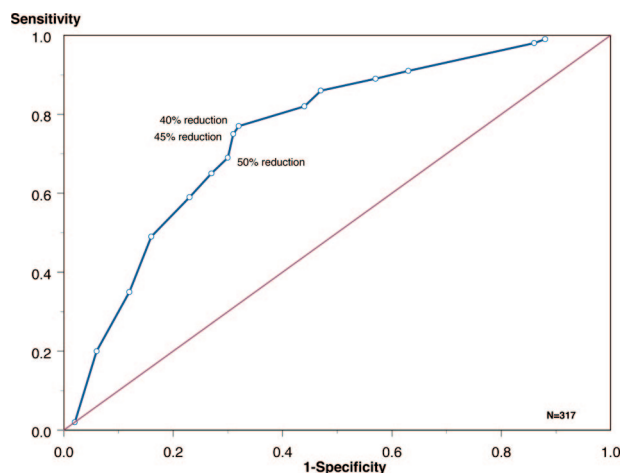


Figure 2. Receiver operator curve relating percentage reduction of dFLC* after 2 months of therapy to overall ECOG response. A total of 40% reduction → sensitivity 77% and specificity 68%; 45% reduction → sensitivity 75% and specificity 69%; and 50% reduction → sensitivity 69% and specificity 70%. FLC was expressed as difference between involved and uninvolved FLC, but the results were similar if absolute FLC used in place of the difference.

Is there a correlation between change of serum immunoglobulin free light chain and of urine M-protein?

We evaluated correlations of serum FLC and 24-hour urine total protein and 24-hour urine M-protein using those patients with “evaluable baseline iFLC” defined as iFLC greater than or equal to 0.05 g/L and an abnormal serum rFLC. After excluding the patients missing a second 24-hour urine, the number of patients with 24-hour urine total protein and 24-hour urine M-protein available for analysis was 148 and 101, respectively (Table 5). The correlation coefficients between percentage change of serum immunoglobulin FLCs and urine M-protein ranged between 0.37 and 0.62. The absence of a strong correlation persisted regardless of whether the change in the absolute values iFLC, the dFLC, or the rFLC was used. An identical analysis was done using the “evaluable baseline iFLC” defined as iFLC greater than or equal to 0.1 g/L, and the results were no different (data not shown).

Because the FLCs are partially dependent on renal function, patients were split into groups with normal and abnormal renal function, but the correlation coefficients comparing changes in serum FLCs to 24-hour urine M-protein continued to run in the same range of 0.27 to 0.69. In addition, the same analyses were performed using the 24-hour urine total protein to account for the potential confounding factor introduced by variable gating of urinary M-proteins; however, the correlation coefficients were no better (Table 5).

Do changes of FLC after 2 cycles of therapy predict for overall or progression-free survival?

Although FLC response after 2 cycles of therapy predicted for eventual overall response, FLC response at 2 months did not

Table 5. Association between percentage change in FLC* and urine protein in patients with abnormal baseline FLC† by baseline renal function status: Spearman correlation coefficient (P)

	Urine M-spike			Urine total protein		
	Overall population	Normal renal function	Abnormal renal function	Overall population	Normal renal function	Abnormal renal function
Urine-only patients plus serum and urine patients, n	101	81	20	148	119	29
iFLC	.60 (<.001)	.67 (<.001)	.55 (.01)	.48 (<.001)	.48 (<.001)	.62 (<.001)
dFLC	.60 (<.001)	.66 (<.001)	.61 (.004)	.46 (<.001)	.47 (<.001)	.55 (.002)
rFLC	.40 (.001)	.43 (<.001)	.27 (.3)	.32 (.003)	.30 (.001)	.42 (.003)
Urine-only patients, n	32	26	6	37	32	5
iFLC	.58 (<.001)	.69 (.001)	.49 (.3)	.51 (.001)	.54 (.002)	.70 (.2)
dFLC	.58 (<.001)	.69 (.001)	.49 (.3)	.51 (.001)	.54 (.001)	.70 (.2)
rFLC	.54 (.002)	.61 (.001)	.31 (.5)	.55 (<.001)	.54 (.001)	.70 (.2)
Serum and urine patients, n	69	55	14	111	87	24
iFLC	.62 (<.001)	.65 (<.001)	.58 (.03)	.49 (<.001)	.48 (<.001)	.62 (.001)
dFLC	.62 (<.001)	.64 (<.001)	.62 (.02)	.49 (<.001)	.46 (<.001)	.53 (.007)
rFLC	.37 (.002)	.38 (.004)	.28 (.3)	.28 (.003)	.25 (.02)	.37 (.08)

Data are correlation coefficients (P) except where shown.

*Percent change = (2-month – baseline)/baseline.

†Abnormal ratio and dFLC ≥ 5 mg/dL.

predict for OS or PFS regardless of which definition for FLC response was applied (data not shown) and whether or not the survival estimates were done as a univariate analysis or a multivariate analysis incorporating other prognostic variables, including treatment arm, plasma cell labeling index, beta-2 microglobulin, and albumin (data not shown). This held true if reductions were based on the 50% reduction rule or reductions by tertiles. Moreover, whether the cohort of patients with baseline iFLC greater than 0.1 g/L was used or the cohort defined by baseline abnormal ratio and dFLC greater than or equal to 0.05 g/L (5 mg/dL) was used, the results were similar, that is, FLC response did not predict for either OS or PFS.

Discussion

We have made several important observations about immunoglobulin FLC in a large, mature dataset of myeloma patients treated with alkylator-based therapy. We explored how best to use the FLC in nonoligosecretory myeloma patients, including: minimal values for FLC to be deemed “evaluable,” best cut-point for FLC response, best use of involved and uninvolved FLC for measuring response, and relationship between the serial measurements of 24-hour urine protein and serum FLC. We also confirmed that elevations in baseline FLC are associated with poorer PFS and OS.^{13,14} Finally, we demonstrated that FLC response after 2 months of therapy is superior to early M-protein measurement to predict overall response but did not predict for OS or PFS.

In both the amyloidosis and the myeloma literature, a baseline value of iFLC of 0.1 g/L has empirically assigned as “measurable” or evaluable disease.^{15,16} We found that, on comparing that cut-point to a cut-point of 0.05 g/L in context of an abnormal FLC ratio, the second definition captured 9% more patients and that the performance characteristics of using early FLC response to predict for eventual overall response were comparable regardless of which definition was used to define the evaluable population. It was gratifying to find that the best predictor for overall response was a 45% to 50% reduction in the dFLC, justifying the 50% reduction empirically chosen by 2 consensus panels.^{15,16} Moreover, we demonstrated that the use of a 50% reduction in dFLC is preferred over a 50% rFLC alone and over a 50% reduction of iFLC contemporaneous with a 50% rFLC. This finding is not surprising.

Although the FLC ratio is an excellent predictor of prognosis in a number of plasma cell disorders^{5,6,8-11} and although it can help differentiate polyclonal rises in FLCs from monoclonal increases,¹ it is skewed by significant drops in the uninvolved FLCs that are induced by chemotherapy.⁶ iFLC or dFLC is preferred over the rFLC for serial measurements.

We also demonstrated that, although there is a correlation between the changes of serum FLC and 24-hour urine M-protein, the degree of correlation found was insufficient to consider the tests interchangeable. One could argue that our negative findings are a function that the “gold standard” (the 24-hour urine M-protein) is an imperfect “gold standard” based on its dependency on patient technique and compliance with a proper 24-hour collection. This is certainly a possibility, but it cannot be proved or disproved by the present study or any study published to date. The burden of evidence must rest on the new test before current 24-hour urine M-protein measurements can be supplanted by the serum FLC measurement. Multiple investigators have shown that there is no correlation between serum FLC and urine M-protein,²¹⁻²³ but there are 2 publications that appear to contradict our findings.^{21,24} Both of these studies are limited by sample size and/or insufficient information regarding details of the analyses.

Finally, we confirmed that early FLC decline was predictive of ultimate response and that reductions in FLCs preceded that of intact immunoglobulin declines²⁵ but in a larger and more systematic fashion. The concept that an early reduction of immunoglobulin light chains is predictive of eventual response is not new. In 1982, McLaughlin and Alexanian found that 24-hour urine M-protein reductions predated intact serum immunoglobulin responses and were predictive of ultimate response.²⁶ They even suggested that “more frequent measurements of Bence Jones protein excretion, such as weekly after the start of chemotherapy, might have confirmed response and resistance trends even earlier in some patients.”²⁶ van Rhee et al have made a contradictory observation in a cohort of 303 patients.¹⁴ They demonstrated that early FLC reduction was predictive of worse PFS and OS. The discrepancy between the van Rhee et al study and ours may be the result of treatments (VBMCP in the present study and VDT-PACE [bortezomib, dexamethasone, thalidomide, cisplatin, doxorubicin, cyclophosphamide, and etoposide] in the Arkansas cohort) and timing of FLC measurement but more likely the result of methodology. As analyzed, the depth of early FLC reduction appears merely

to have been a surrogate for high baseline FLC, which has consistently been shown to be an adverse prognostic finding.^{13,14} After 2 cycles of VDT-PACE,¹⁴ their FLC reduction tertiles were less than 75%, 75% to less than 96%, and 96% to 100%, with respective 24-month estimated survival rates of 91%, 93%, and 79%; however, if one does the math, none of the patients in the lowest baseline tertile and an undefined number in the second baseline tertile would be eligible to have a 96% reduction in iFLC without dropping below pathologic levels of iFLC. Even in their reanalysis of the data, the problem was not fully addressed because baseline iFLC was not included in the multivariate analysis that evaluated the role of depth of FLC response.²⁷ Therefore, according to their methodology, patients with low baseline iFLC (best prognosis patients) were destined to be "low responders."²⁸

So what role does serum FLC measurement play in patients with disease measurable by electrophoresis? Should the FLC assay be run each cycle to foreshadow ultimate response? Our study does not definitively answer these questions. The sensitivity and specificity of early FLC response for ultimate response were 61% and 87%, respectively, but response did not translate into improved PFS or OS. Historically, lack of progression rather than paraprotein response has been most predictive of OS.²⁹ The generalizability of early free light response demonstrated in our study may also be limited by the fact that VBMCP-like therapy is obsolete as induction therapy and that our analysis was restricted to only one time point after therapy. Our data do not inform on the value of serial FLC measurement to predict for early relapse. Although it has been shown that rising FLC predicts relapse sooner than do standard electrophoretic measurements,³⁰⁻³² the utility of this observation is uncertain because there are no data lending credence to the concept that earlier detection or treatment of relapse leads to better outcome.

In conclusion, it appears that the primary utility of immunoglobulin FLC in newly diagnosed myeloma patients with disease measurable by electrophoresis is of baseline prognostic value. Serial measurements in this cohort do not appear to have added value. If serial measurements are done, among changes in the absolute value of iFLC, the dFLC, and the rFLC, changes of the rFLC are least useful.

Acknowledgments

This work was supported in part by Eastern Cooperative Oncology Group (Robert L. Comis, Chair) and supported in part by Public Health Service (grants CA62242, U10CA21115, and the Robert A. Kyle Hematologic Malignancies Fund). Reagents were supplied from the Binding Site LTD.

Authorship

Contribution: A.D. designed research, performed research, analyzed data, and wrote the paper; L.Z. analyzed data; E.B. designed research; J.A.K., M.S., R.D., K.H., A.R.B., R.A.K., M.M.O., and P.R.G. performed research and analyzed data.

Conflict-of-interest disclosure: A.R.B. is an employee of the Binding Site LTD. J.A.K. receives reagents from the Binding Site LTD. The remaining authors declare no competing financial interests.

Correspondence: Angela Dispenziera, Mayo Clinic, 200 First Street SW, Rochester, MN 55905; e-mail: dispenziera.angela@mayo.edu.

References

- 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.
- Drayson M, Tang LX, Drew R, Mead GP, Carr-Smith H, Bradwell AR. Serum free light-chain measurements for identifying and monitoring patients with nonsecretory multiple myeloma. *Blood*. 2001;97:2900-2902.
- Abraham RS, Katzmann JA, Clark RJ, Bradwell AR, Kyle RA, Gertz MA. Quantitative analysis of serum free light chains: a new marker for the diagnostic evaluation of primary systemic amyloidosis. *Am J Clin Pathol*. 2003;119:274-278.
- Katzmann JA, Abraham RS, Dispenziera A, Lust JA, Kyle RA. Diagnostic performance of quantitative kappa and lambda free light chain assays in clinical practice. *Clin Chem*. 2005;51:878-881.
- Lachmann HJ, Gallimore R, Gillmore JD, et al. Outcome in systemic AL amyloidosis in relation to changes in concentration of circulating free immunoglobulin light chains following chemotherapy. *Br J Haematol*. 2003;122:78-84.
- Dispenziera A, Lacy MQ, Katzmann JA, et al. Absolute values of immunoglobulin free light chains are prognostic in patients with primary systemic amyloidosis undergoing peripheral blood stem cell transplantation. *Blood*. 2006;107:3378-3383.
- Wolff F, Thyry C, Willems D. Assessment of the analytical performance and the sensitivity of serum free light chains immunoassay in patients with monoclonal gammopathy. *Clin Biochem*. 2007;40:351-354.
- Rajkumar SV, Kyle RA, Therneau TM, et al. Presence of monoclonal free light chains in the serum predicts risk of progression in monoclonal gammopathy of undetermined significance. *Br J Haematol*. 2004;127:308-310.
- Rajkumar SV, Kyle RA, Therneau TM, et al. Serum free light chain ratio is an independent risk factor for progression in monoclonal gammopathy of undetermined significance. *Blood*. 2005;106:812-817.
- Dingli D, Kyle RA, Rajkumar SV, et al. Immunoglobulin free light chains and solitary plasmacytoma of bone. *Blood*. 2006;108:1979-1983.
- Dispenziera A, Kyle RA, Katzmann JA, et al. Immunoglobulin free light chain ratio is an independent risk factor for progression of smoldering (asymptomatic) multiple myeloma. *Blood*. 2008;111:785-789.
- Bradwell AR, Carr-Smith HD, Mead GP, Harvey TC, Drayson MT. Serum test for assessment of patients with Bence Jones myeloma. *Lancet*. 2003;361:489-491.
- Kyrtonis MC, Vassiliakopoulos TP, Kafasi N, et al. Prognostic value of serum free light chain ratio at diagnosis in multiple myeloma. *Br J Haematol*. 2007;137:240-243.
- 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.
- Gertz MA, Comenzo R, Falk RH, et al. Definition of organ involvement and treatment response in immunoglobulin light chain amyloidosis (AL): a consensus opinion from the 10th International Symposium on Amyloid and Amyloidosis, Tours, France, 18-22 April 2004. *Am J Hematol*. 2005;79:319-328.
- Durie BG, Harousseau JL, Miguel JS, et al. International uniform response criteria for multiple myeloma. *Leukemia*. 2006;20:1467-1473.
- Oken MM, Leong T, Lenhard RE Jr, et al. The addition of interferon or high dose cyclophosphamide to standard chemotherapy in the treatment of patients with multiple myeloma: phase III Eastern Cooperative Oncology Group Clinical Trial EST 9486. *Cancer*. 1999;86:957-968.
- 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.
- Bloch DA. Comparing two diagnostic tests against the same "gold standard" in the same sample. *Biometrics*. 1997;53:73-85.
- Greipp PR, San Miguel J, Durie BG, et al. International staging system for multiple myeloma. *J Clin Oncol*. 2005;23:3412-3420.
- Abraham RS, Clark RJ, Bryant SC, et al. Correlation of serum immunoglobulin free light chain quantification with urinary Bence Jones protein in light chain myeloma. *Clin Chem*. 2002;48:655-657.
- Nowrousian MR, Brandhorst D, Sammet C, et al. Serum free light chain analysis and urine immunofixation electrophoresis in patients with multiple myeloma. *Clin Cancer Res*. 2005;11:8706-8714.
- Singhal S, Stein R, Vickrey E, Mehta J. The serum-free light chain assay cannot replace

- 24-hour urine protein estimation in patients with plasma cell dyscrasias. *Blood*. 2007;109:3611-3612.
24. Alyanakian MA, Abbas A, Delarue R, Arnulf B, Aucouturier P. Free immunoglobulin light-chain serum levels in the follow-up of patients with monoclonal gammopathies: correlation with 24-hr urinary light-chain excretion. *Am J Hematol*. 2004;75:246-248.
25. Mead GP, Carr-Smith HD, Drayson MT, Morgan GJ, Child JA, Bradwell AR. Serum free light chains for monitoring multiple myeloma. *Br J Haematol*. 2004;126:348-354.
26. McLaughlin P, Alexanian R. Myeloma protein kinetics following chemotherapy. *Blood*. 1982;60:851-855.
27. van Rhee F, Crowley J, Barlogie B. Response: Top tertile SFLC reduction indeed is an independent feature of myeloma aggressiveness. *Blood*. 2008;111:2491.
28. Dispenzieri A. Is early, deep free light chain response really an adverse prognostic factor? *Blood*. 2008;111:2490-2491.
29. Durie BG, Jacobson J, Barlogie B, Crowley J. Magnitude of response with myeloma frontline therapy does not predict outcome: importance of time to progression in southwest oncology group chemotherapy trials. *J Clin Oncol*. 2004;22:1857-1863.
30. Tate JR, Mollee P, Dimeski G, Carter AC, Gill D. Analytical performance of serum free light-chain assay during monitoring of patients with monoclonal light-chain diseases. *Clin Chim Acta*. 2007;376:30-36.
31. Mayo MM, Johns GS. Serum free light chains in the diagnosis and monitoring of patients with plasma cell dyscrasias. *Contrib Nephrol*. 2007;153:44-65.
32. Drayson M, Begum G, Basu S, et al. Effects of paraprotein heavy and light chain types and free light chain load on survival in myeloma: an analysis of patients receiving conventional-dose chemotherapy in Medical Research Council UK multiple myeloma trials. *Blood*. 2006;108:2013-2019.