International Staging System for Multiple Myeloma


A B S T R A C T

Purpose
There is a need for a simple, reliable staging system for multiple myeloma that can be applied internationally for patient classification and stratification.

Patients and Methods
Clinical and laboratory data were gathered on 10,750 previously untreated symptomatic myeloma patients from 17 institutions, including sites in North America, Europe, and Asia. Potential prognostic factors were evaluated by univariate and multivariate techniques. Three modeling approaches were then explored to develop a staging system including two nontree and one tree survival assessment methodologies.

Results
Serum beta2-microglobulin (SB2-M), serum albumin, platelet count, serum creatinine, and age emerged as powerful predictors of survival and were then used in the tree analysis approach. A combination of SB2-M and serum albumin provided the simplest, most powerful and reproducible three-stage classification. This new International Staging System (ISS) was validated in the remaining patients and consists of the following stages: stage I, SB2-M less than 3.5 mg/L plus serum albumin ≥ 3.5 g/dL (median survival, 62 months); stage II, neither stage I nor III (median survival, 44 months); and stage III, SB2-M ≥ 5.5 mg/L (median survival, 29 months). The ISS system was further validated by demonstrating effectiveness in patients in North America, Europe, and Asia; in patients less than and ≥ 65 years of age; in patients with standard therapy or autotransplantation; and in comparison with the Durie/Salmon staging system.

Conclusion
The new ISS is simple, based on easy to use variables (SB2-M and serum albumin), and recommended for early adoption and widespread use.

J Clin Oncol 23. © 2005 by American Society of Clinical Oncology

INTRODUCTION

The outcome for patients with multiple myeloma is highly variable. Although the median overall survival time is 3 to 4 years, the range is from less than 6 months to greater than 10 years. This variability derives from heterogeneity in both myeloma cell biology and multiple host factors. Knowledge of tumor and host factors associated with prognosis is critical for understanding disease outcome, identifying risk groups, and optimizing patient treatment.

Studies conducted in the 1960s and early 1970s identified a number of clinical and laboratory parameters that are independent predictors of survival duration including hemoglobin level, serum calcium, serum creatinine, and severity of bone lesions.1,2 Subsequently, combinations of prognostic factors were suggested for staging classification of myeloma patients.3-5 In 1975, Durie and Salmon6 introduced a staging system, the Durie/Salmon (DS) system, using commonly available clinical parameters that predicted...
myeloma cell tumor burden. Factors in the DS classification included the level and type of monoclonal protein, hemoglobin, calcium level, and number of bone lesions. Creatinine level (substage A: serum creatinine < 2 mg/dL; and substage B: serum creatinine ≥ 2 mg/dL) further defined lower versus higher risk patients in each of the three tumor mass stages. The DS system was widely adopted as the standard for prognostication in myeloma. The number of lytic lesions on routine radiographs (skeletal survey), an important element of the DS system, was widely adopted as the standard for prognostication in myeloma. The number of lytic lesions on routine radiographs (skeletal survey), an important element of the DS system, is unfortunately observer dependent. In an effort to ensure a more objective pretreatment classification of patients, several staging methods were proposed.7–9 In the 1980s, serum beta2-microglobulin (Sβ2M) emerged as the single most powerful prognostic factor and was considered a simple reliable predictor of survival duration.10–13

Subsequently, other prognostic factors were introduced, including serum levels of C-reactive protein,14 albumin,7 and the proliferative activity of bone marrow plasma cells assessed by labeling index15 or flow cytometry cell-cycle (S phase) analysis.16 Combining these factors with Sβ2M provided improved prognostic stratification compared with Sβ2M alone.7,14–16 However, there was no consensus about which factors should be combined with Sβ2M, and there was no consensus on cutoff values for Sβ2M or other variables.

Recently, conventional cytogenetics by karyotyping has emerged as a relevant prognostic factor in myeloma patients. Deletion of chromosome 13 (del 13) is the most common and the most significant prognostic abnormality observed.17–20 Conventional cytogenetics is able to identify abnormalities in the myeloma clone in 20% to 30% of patients. Fluorescent in situ hybridization (FISH) techniques may offer more sensitive and specific identification of such critical abnormalities. Although prognostically important correlations have emerged, practical application of these techniques has been hampered by lack of standardization, costs, and restricted availability. The favorable experience of international cooperative efforts for the design of a prognostic index in non-Hodgkin’s lymphoma21 led us to embark on a project to design a staging system for multiple myeloma that would be based on widely available, objective parameters used around the world.

### RESULTS

#### Patient Characteristics

Median age of the 10,750 patients was 60 years; 57% of patients were male, and 60% had IgG isotype, 24% had IgA isotype, 11% had light-chain isotype only, 3% had IgD isotype, and 2% had biclonal or other isotype. Median serum M-protein level was 3.9 g/dL, hemoglobin level was 10.5 g/dL,
platelet level was 222 × 100/µL, creatinine level was 1.1 mg/dL, Sβ2M level was 3.8 µg/mL, albumin level was 3.6 g/dL, and bone marrow plasma cell percent was 40.0%. Forty-three percent of patients had three or more bone lesions, and 25% had pathologic fractures. Overall survival time for the entire group was 44 months.

**Preliminary Prognostic Factor Analysis**

In preparation for developing a staging system, half of the total myeloma patients (5,383 patients) were randomly selected as a training sample.

**Univariate and Multivariate Survival Analysis**

Table 1 allows comparison of variables in both univariate and multivariate models and lists the 10 most important prognostic factors in univariate analyses. The variables are ranked by hazard ratio, with all being significant at the P ≤ .001 level. The numbers and percentages of patients in each risk category are also listed. Sβ2M and serum albumin were the most consistent, broadly applicable, prognostic factors correlated with survival duration. Attention is drawn to low platelet count (ranked 2) and high serum creatinine (ranked 4) results; although these parameters rank highly, they identify small patient subsets of 12% and 17%, respectively. Conversely, high Sβ2M (ranked 1) and low serum albumin (ranked 8) identify larger patient subsets of 56% and 40%, respectively. Other relevant prognostic factors were age, hemoglobin, calcium, lactate dehydrogenase, and bone marrow plasma cell infiltration (Table 1). Additional, although weaker, prognostic factors (ranked 11 to 15) on univariate analysis were C-reactive protein, Ig isotype, size of M-component, and extent of bone lesions (data not shown). There were 5,894 patients with IgG isotype, 2,375 with IgA isotype, and 1,035 with light-chain only isotype. The median survival time for the IgG patients was longer (49 months) compared with the IgA patients (40 months) and light-chain patients (35 months). The P values are as follows: IgG versus IgA, P < .001; IgG versus light chain only, P < .001; and IgA versus light chain only, P < .009. Regarding sex, there were 4,597 female patients and 6,153 male patients with median survival times of 45 and 44 months, respectively (slight advantage for female patients).

**Development of a Myeloma Staging System**

The information from univariate and multivariate analyses was used to explore three modeling approaches. The most significant prognostic factors were assessed using the following three methods: (1) the weighted variable model; (2) the model based on the number of risk factors occurring in an individual patient; and (3) the survival tree model in which risk factors present at each branch point are sequentially reassessed.

The weighted variable model was derived from the Cox regression approach using continuous variables. The regression equation was used to derive a prediction of risk, which was then stratified by tertiles to form risk groups. Similarly, for a model based on the number of risk factors, the risk groups were based on the five factors identified in Table 1 (right side) as the major prognostic factors in a Cox multivariate regression analysis using, in this case, dichotomous variables. Sβ2M and serum albumin were the dominant independent prognostic factors in all three models.

With the survival tree model, a three-stage system using Sβ2M and albumin provided the most highly statistically significant results. From now on, this system is called ISS (Table 2). Median survivals were as follows: stage I, 62 months; stage II, 45 months; and stage III, 29 months (P < .0001 for differences). Patient numbers were well distributed across the three stages (stage I, 28.9%; stage II, 37.5%; and stage III, 33.6%). Among the three methods for developing a staging system, we chose the survival tree approach as being the simplest, most effective and readily understood method.

### Table 1. Comparison of Univariate and Multivariate Correlates of Survival Duration*

<table>
<thead>
<tr>
<th>No. of Patients/Total No.</th>
<th>%</th>
<th>Hazard Ratio</th>
<th>Variables</th>
<th>Hazard Ratio</th>
<th>Sequence of Entry</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,428/4,313</td>
<td>56</td>
<td>1.81</td>
<td>Sβ2M ≥ 3.5 mg/dL</td>
<td>1.81</td>
<td>Sβ2M 1</td>
</tr>
<tr>
<td>570/4,878</td>
<td>12</td>
<td>1.73</td>
<td>Platelet count (Platelets) &lt; 130,000/µL</td>
<td>1.63</td>
<td>Platelets 2</td>
</tr>
<tr>
<td>1,842/5,358</td>
<td>34</td>
<td>1.67</td>
<td>Age ≥ 65 years</td>
<td>1.28</td>
<td>CALC 4</td>
</tr>
<tr>
<td>868/5,181</td>
<td>17</td>
<td>1.66</td>
<td>Serum CREAT ≥ 2 mg/dL</td>
<td>1.28</td>
<td>CREAT 5</td>
</tr>
<tr>
<td>533/2,050</td>
<td>26</td>
<td>1.5</td>
<td>Serum LDH value &gt; normal</td>
<td>1.28</td>
<td>ALB 3</td>
</tr>
<tr>
<td>2,077/5,175</td>
<td>40</td>
<td>1.49</td>
<td>Hemoglobin &lt; 10 g/dL</td>
<td>1.28</td>
<td>Bone marrow plasma cells ≥ 33%</td>
</tr>
<tr>
<td>938/3,100</td>
<td>19</td>
<td>1.44</td>
<td>Performance status &gt; 3</td>
<td>1.28</td>
<td></td>
</tr>
<tr>
<td>1,940/4,770</td>
<td>40</td>
<td>1.4</td>
<td>Serum ALB &lt; 3.5 g/dL</td>
<td>1.28</td>
<td></td>
</tr>
<tr>
<td>1,588/4,754</td>
<td>33</td>
<td>1.32</td>
<td>Serum CALC &gt; 10 mg/dL</td>
<td>1.28</td>
<td></td>
</tr>
<tr>
<td>2,897/4,996</td>
<td>58</td>
<td>1.29</td>
<td>Bone marrow plasma cells ≥ 33%</td>
<td>1.28</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: Sβ2M, serum beta₂-microglobulin; ALB, albumin; CREAT, creatinine; LDH, lactate dehydrogenase; CALC, calcium.

*See text for details.

†These are results of stepwise multivariate regression analysis. Each sequential hazard ratio reflects adjustment for prior variables.
Validation of the New Staging System (ISS) in the Remaining Half of Patients (Validation Sample)

The ISS system was applied to the remaining patients, who constituted half of the whole sample. The discrimination and survival durations were almost identical (stage I, 62 months; stage II, 44 months; and stage III, 29 months) to the training sample (Fig 1 and Table 2).

Because the training and validation samples had almost identical outcomes, the two patient populations were combined into a single group for further analysis. It is important to note that serum albumin added consistent prognostic discrimination versus $s_{\beta_2} M$ alone. Thus, of the 3,157 patients with a low $s_{\beta_2} M$ less than 3.5 mg/dL, 1,020 (32% of these patients; 12.5% of the total population) were classified as stage II because of a low albumin less than 3.5 mg/L. On analyzing the characteristics of patients according to the new ISS, we observed that the more advanced the stage, the higher was the proportion of patients with advanced age, anemia (hemoglobin < 10 g/dL), thrombocytopenia (platelet count < 130,000/μL), high bone marrow infiltration, and poor performance status (Table 3). The frequency of patients with advanced DS stage III (A or B) progressively increased from stage ISS I (38%) to stage ISS II (54%) and stage ISS III (70%) categories. For DS stage IIIB, all patients (100%) were ISS stage III.

As an adjunct to the new staging system, a search was undertaken to identify simple predictors of very poor prognosis (eg, overall survival < 12 months). A type of forward stepwise regression analysis was performed to identify factors associated with very poor risk, as noted in Patients and Methods.28 This technique is called extreme regression. Using this technique, four factors emerged as being helpful in identifying very poor risk patients; these factors were $s_{\beta_2} M$ more than 10 mg/L, serum creatinine more than 4 mg/dL, serum albumin less than 2.5 g/dL, and platelet count less than 130,000/μL. It was possible to fit several different models with this regression technique. With all of the models, the very poor risk group was approximately 5% of the total population. Although significant $P$ values of less than .0001 were obtained, it must be noted that the worst median overall survival time for a poor risk group was 17 months. Thus, there was somewhat limited ability to accurately predict very poor survival with these routine test parameters. Half of the patients had a survival time of more than 17 months and did not fall into the less than 12 months very poor risk group that was being sought.

Cytogenetic data (Table 3) were available in a subset of 390 patients. No strong correlations with stage were observed. The translocation t(4;14) occurred with a lower incidence ($P = .035$) in stage I patients than in stage II and III patients (6% vs 16% and 11%, respectively). Although there was a slight trend for less frequent complex karyotypic abnormalities, deletion 13 by FISH and deletion 13 by cytogenetics in stage I disease, these trends were not statistically significant ($P = .075$ and $P = .162$, respectively). Considering the impact of cytogenetic abnormalities overall, patients with and without cytogenetic abnormalities of any type were compared. The median overall survival for the 113 patients with cytogenetic abnormalities was 42 months vs 69 months for the 277 patients with no cytogenetic abnormalities. The $P$ value for no cytogenetic abnormalities versus cytogenetic abnormalities is $P = .03$.

Other Assessments of the New ISS System

Geographic region. In proposing an international system, it was important to validate the system by geographic region. As illustrated in Figure 2, there was comparable utility in patients from North America, Europe, and Asia. Discriminatory efficacy was also excellent comparing individual institutions versus cooperative groups (data not shown).
Age. Age is not only an important prognostic factor, but it also critically influences treatment options, such as high-dose therapy. Accordingly, we wanted to analyze whether the ISS system applies equally to young and older patients. Thus, although older patients (eg, > 65 years) have poorer survival than younger patients, it is important to note, as illustrated in Figure 3, that the ISS system applies to both groups.

Treatment type. As far as treatment type is concerned, 7,920 patients were treated with standard-dose therapy as the primary modality, whereas 2,807 patients received high-dose therapy with autologous marrow or stem-cell rescue on an intent-to-treat basis (defined as within 9 months of start of therapy). Again, the ISS system discriminated similarly for the two groups, as shown in Figure 4.

Comparisons with the DS system. The survival duration comparisons of the DS system versus the ISS system are listed in Table 4 and Figure 5. Compared with the DS classification, the ISS provides more equal distribution of patients across the three stages. DS stage I patients are underrepresented in these data sets. The 8% value is lower than the typical 20% level, most likely because DS stage I patients are excluded from many protocols involved in the data sets analyzed. Nonetheless, the survival of ISS stage I corresponds exactly to the DS stage IA patients who are incorporated in these analyses. Both groups of patients have median survival times of 62 months. Interestingly, DS stage IIA reflects a similar patient population, with a median survival time of 58 months. ISS stage II patients correspond to DS stage IIIA patients, with median survival times of 44 and 45 months, respectively. Of particular note, ISS stage III identifies DS substage B (serum creatinine ≥ 2 mg/dL) for DS stages I, II, and III (ie, the poor-risk B subset), irrespective of tumor burden. It is obviously helpful to have such patients categorized collectively in ISS stage III. In Table 3, it can be seen that of the 1,382 total patients with a serum creatinine ≥ 2 mg/dL, 82% had a S2M value of more than 5.5 mg/L and, therefore, were classified as ISS stage III. Importantly, the S2M values (ranked number 1 in both univariate and multivariate analyses) are much more powerful versus serum creatinine values (ranked number 4 in univariate and number 5 in multivariate analyses; Table 1).

**DISCUSSION**

The DS clinical staging system has remained the most widely used staging system for over 25 years. Although a few prognostic parameters, such as S2M, have emerged as better predictors of survival duration, there has been no consensus as to the optimal use of single or multiple prognostic factors. The current large international data set of patients with symptomatic myeloma offers the
opportunity to establish a statistically superior and widely accepted new staging system.

The new ISS (Table 2) was developed using univariate and multivariate analyses (Table 1) and three types of modeling approaches. $S\beta_2M$ and serum albumin were selected from the various potential prognostic factors both because of the statistical power in various models as well as the known wide availability of these two simple inexpensive laboratory tests. The inclusion of serum albumin as the second parameter added significantly in defining 1,020 patients (12.5% of total population), now identified as ISS stage II, who would otherwise have been classified as stage I based on low ($< 3.5 \text{ mg/L}$) $S\beta_2M$ alone. The large data set afforded the opportunity to establish clear cutoff values to identify the three stages in the new ISS system (summarized in Table 2 and displayed in Fig 5). The survival differences were reproducibly demonstrated in the test and validation datasets. The broad applicability of the ISS system was further illustrated with validation by geographic area, patient age, and treatment type and in comparison with the DS staging system (Figs 1 and 2, Table 4). Of particular note, ISS stage III is clearly delineated as a poor-risk group (39% of patients), with a median survival time of 29 months (Table 4) versus the more mixed and numerous DS stage IIIA (49% of patients; median survival time, 45 months) and stage IIIB (17% of patients; median survival time, 24 months).

Early attempts to improve on the DS staging system were not widely adopted.\textsuperscript{7} But now, $S\beta_2M$ is widely recognized as the single most important variable predicting survival.\textsuperscript{13} When added to $S\beta_2M$, serum albumin level was known to add significantly to prognostication.\textsuperscript{7} There was much debate as to whether these were sufficient prognostic factors or whether better prognostic factors were required.\textsuperscript{29-33} However, in the absence of any additional, powerful prognostic factors, further analyses using $S\beta_2M$ and serum albumin were conducted. This led to a $S\beta_2M$ and serum albumin staging system\textsuperscript{29} developed by the Southwest Oncology Group. The newly developed and proposed ISS system thus extends and validates these prior observations.

The following question emerges: why are $S\beta_2M$ and serum albumin such powerful prognostic factors? $S\beta_2M$ reflects not only tumor mass and renal function but also other as yet unknown parameters, possibly including...
immune function.7,34,35 The specific cause of decreased albumin in some multiple myeloma patients is not certain; however, a lower albumin may reflect effects on the liver by interleukin-6 produced by the microenvironment of myeloma cells.34,35 The strong correlations between serum levels of $\beta_2$M and serum albumin and myeloma patient survival imply connections to important underlying mechanisms. There are several clues in the published literature,36-41 but to date, the underlying biology remains to be explored.

The ISS provides useful prognostic groupings in a variety of situations (in patients aged greater or less than 65 years, Fig 3; with conventional or high-dose transplantation therapy, Fig 4; in Europe, Asia, and North America, Fig 2; and in single institutions or cooperative groups, data not displayed). Because the levels of $\beta_2$M and albumin are now specified by the ISS, it is critical that laboratory variation be minimized by standardizing methods used to determine their levels, specifically in multiple myeloma. That work and standardization is underway by the Nordic Myeloma Study Group, who are members of the International Myeloma Working Group.

We conclude that the ISS staging system is broadly useful and that it will provide a sound base for more advanced studies in the future. Identification of highest risk patients was achieved in only a small number of patients (5% to 9%) using standard variables. Better identification of such patients may require a more refined cytogenetic and molecular genetic classification. As more data and follow-up research are collected, the ISS system may continue to evolve. However, even in its current state, the ISS provides valuable insights into the prognosis of multiple myeloma patients.

### Table 4. Comparison Between Durie/Salmon and ISS Staging Systems: Survival Duration by Stage in Months

<table>
<thead>
<tr>
<th>Stage</th>
<th>Durie/Salmon</th>
<th>ISS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% of Patients*</td>
<td>Median Survival (months)</td>
</tr>
<tr>
<td>IA</td>
<td>7.5</td>
<td>62</td>
</tr>
<tr>
<td>IB</td>
<td>0.5</td>
<td>22</td>
</tr>
<tr>
<td>IIA</td>
<td>22</td>
<td>58</td>
</tr>
<tr>
<td>IIB</td>
<td>4</td>
<td>34</td>
</tr>
<tr>
<td>IIIA</td>
<td>49</td>
<td>45</td>
</tr>
<tr>
<td>IIIIB</td>
<td>17</td>
<td>24</td>
</tr>
</tbody>
</table>

Abbreviation: ISS, International Staging System.
*Percentage of patients falling into each staging category.
become available, the International Myeloma Working Group plans to develop a second staging system using conventional and FISH cytogenetics, molecular genetics, proteomics, and S-phase analysis for use by reference centers and eventually for all patients with myeloma.

Appendix

The journal of Clinical Oncology limits author lists to 20. The authors listed for this manuscript are therefore the members of the International Myeloma Working Group who represent institutions and groups contributing patient data for the analyses. Meral Beksaç, Ankara University, Ibnî Sina Hospital, Ege and İnönü Universities, Ankara and Istanbul, Turkey also contributed patient data. Other members of the International Myeloma Working Group who contributed to this study include the following: Raymond Alexanian, University of Texas, Houston, TX; Kenneth Anderson, Dana-Farber Cancer Center, Boston, MA; Michel Attali, Institut de Biologie, Nantes, France; Hervé Avet-Loiseau, Institut de Biologie, Nantes, France; Ismet Aydogdu, İnönü University, Turkey; Regis Bataille, University of Nantes, France; William Bensinger, Fred Hutchinson Cancer Research Center, Seattle, WA; Peter Bergsagel, Cornell Medical Center, New York, NY; Seckin Cagirgan, Ege University, Turkey; Michele Cavo, Istituto di Ematologia, University of Bologna, Bologna, Italy; Ray Comenzo, Sloan-Kettering Cancer Research Center, New York, NY; William Dalton, University of South Florida, Tampa, FL; Meletios Dimopoulos, University of Athens, Athens, Greece; Mark Drayson, University of Birmingham, Birmingham, United Kingdom; Thierry Facon, Institut de Biologie, Nantes, France; Dorotea Fantl, Istituto di Biologia, Nantes, France; Bataille R, Durie BGM, Grenier J, et al: Serum beta 2 microglobulin, a tumor marker of lymphoproliferative disorders. Lancet 2:108-109, 1978


22. Durie BGM, Salmon SE. Recent Advances in Haematology. Edinburgh, United Kingdom, Churchill Livingstone, 1977


