The Prognostic Significance of Serum β2 Microglobulin Levels in Acute Myeloid Leukemia and Prognostic Scores Predicting Survival: Analysis of 1,180 Patients

Apostolia-Maria Tsimberidou,1 Hagop M. Kantarjian,1 Sijin Wen,2 Susan O’Brien,1 Jorge Cortes,1 William G. Wierda,1 Charles Koller,1 Sherry Pierce,1 Mark Brandt,1 Emil J. Freireich,1 Michael J. Keating,1 and Elihu H. Estey1

Abstract

Purpose: Serum β2 microglobulin (β2M) is prognostic in other hematologic malignancies; therefore, we evaluated its prognostic significance in acute myeloid leukemia (AML).

Experimental Design: Multivariate analyses were used to examine the effect of pretreatment serum β2M levels on clinical outcomes in patients with AML. β2M was associated with poorer survival in older but not younger patients. We thus fit separate Cox survival models in patients above and below age 60 years treated with remission induction therapy containing high-dose cytarabine (n = 1,280). In each age group, 50% of the patients were used to develop the model, which was tested in the other 50%. Resampling methods were also used to validate the independent prognostic significance of covariates.

Results: In patients 60 years or older (n = 591), poorer risk cytogenetics; poorer performance status; and higher levels of β2M, uric acid, and lactate dehydrogenase were each found to independently predict shorter survival and formed the basis of a scoring system. A similar approach was used in patients younger than 60 years (n = 589), with poorer risk cytogenetics, poorer performance status, older age, higher hemoglobin level, and higher leukocyte count predicting a shorter survival and forming the basis of the scoring system. Higher β2M levels were an adverse independent factor for response, survival, relapse-free survival, and event-free survival in older but not in younger patients.

Conclusions: Serum β2M levels can help predict outcome in patients ≥60 years with untreated AML, and their use is strongly encouraged.

β2 microglobulin (β2M) is a 12-kDa light chain protein that binds to the α chain of class I MHC (MHC1) molecules [e.g., human leukocyte antigen (HLA)-A, HLA-B, and HLA-C; ref. 1]. β2M has been described as playing a dual role in MHC1 molecules. It is a structural subunit of the assembled complex. It is also a chaperone, with a direct effect on the folding of the MHC1 heavy chain by facilitating the interaction of MHC1 heavy chain with other chaperones, such as calreticulin, tapasin, transporters associated with antigen processing, and others (2). With physiologic concentrations of high-affinity peptides or any concentration of lower-affinity peptides, β2M levels limit the folding of MHC1 molecules (3). In the absence of β2M, most MHC1 molecules are not expressed efficiently on the surface of cells (4, 5). Upon metabolism and degradation of HLA, β2M is dissociated from the heavy chain and is released in its free form into the extracellular fluids (6). β2M is filtered almost exclusively by the glomerulus and is most efficiently reabsorbed by the cells of the proximal tubules under diverse physiologic conditions (6).

Serum β2M levels are known to reflect renal function and membrane turnover, which is associated with tumor mass and growth rate (7–10). Elevated serum β2M levels are reported to predict poor survival in several hematologic malignancies, which include multiple myeloma (11), low-grade lymphomas (12), large-cell lymphomas (13–15), Hodgkin’s lymphoma (16, 17), acute lymphoblastic leukemia (18), Philadelphia chromosome-positive chronic myeloid leukemia (19), chronic lymphocytic leukemia (20, 21), and myelodysplastic syndromes (22, 23).

The prognosis of patients with acute myeloid leukemia (AML) varies. Older age, poor risk cytogenetics, and performance status (24–35) are most commonly used to predict clinical outcomes. However, prognostic heterogeneity still exists and novel prognostic factors are being developed. Although some reports on the prognostic significance of β2M in AML have been published (23, 36, 37), the independent prognostic
role of $\beta_2$M in AML has not been established. Here, we examine whether $\beta_2$M level is an independent prognostic factor in untreated AML that could be added to known prognostic factors to reduce prognostic variation.

**Patients and Methods**

**Patients.** We searched the AML database for patients who presented to The University of Texas M. D. Anderson Cancer Center with newly diagnosed AML ($\geq$20% myeloblasts) from 1990 through 2005. This database includes consecutive patients with AML or MDS seen at M. D. Anderson in the Department of Leukemia since 1985. Patients previously classified with refractory anemia with excess blasts in transformation were reclassified as AML. A total of 2,014 patients were identified, and pretreatment levels of $\beta_2$M were available in 64% (i.e., 1,293 patients). Serum $\beta_2$M levels were quantified by RIA (Pharmacia $\beta_2$ Micro Ria; Pharmacia Diagnostic; reference range, 0.7-2.0 mg/L). Treatment for AML varied during the 16 years depicted here, and for convenience we divided the patients into those who were given 1-$h_2$-D-arabinofuranosylcytosine (ara-C) and those who were not. All patients included in the prognostic models received remission induction therapy with high-dose ara-C, defined as >0.5 g daily for 3 to 6 days for patients younger than 60 years and for 2 to 3 days for patients 60 years or older. Responders were to receive high-dose ara-C as maintenance therapy. All protocol patients gave informed consent. The study was approved by the M. D. Anderson Cancer Center Institutional Review Board and was conducted in accordance with the Declaration of Helsinki.

### Table 1. Association between $\beta_2$M and major patient characteristics in 1,293 patients with $\beta_2$M measurements and in patients $\geq 60$ y

<table>
<thead>
<tr>
<th>Variable</th>
<th>n = 1,293 (any age)</th>
<th>n = 703 (age $\geq 60$ y)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\beta_2$M (mg/L)</td>
<td>$\beta_2$M (mg/L)</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>Mean</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>552</td>
<td>3.10</td>
</tr>
<tr>
<td>Male</td>
<td>741</td>
<td>3.91</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>African-American</td>
<td>93</td>
<td>3.79</td>
</tr>
<tr>
<td>Caucasian</td>
<td>1,041</td>
<td>3.58</td>
</tr>
<tr>
<td>Hispanic</td>
<td>128</td>
<td>3.43</td>
</tr>
<tr>
<td>Other</td>
<td>31</td>
<td>2.93</td>
</tr>
<tr>
<td>Cytoegenetics, risk</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Favorable</td>
<td>87</td>
<td>2.93</td>
</tr>
<tr>
<td>Intermediate</td>
<td>781</td>
<td>3.68</td>
</tr>
<tr>
<td>Poor</td>
<td>374</td>
<td>3.56</td>
</tr>
<tr>
<td>Insufficient metaphases</td>
<td>51</td>
<td>2.93</td>
</tr>
<tr>
<td>RAS mutation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>228</td>
<td>3.19</td>
</tr>
<tr>
<td>Yes</td>
<td>44</td>
<td>4.19</td>
</tr>
<tr>
<td>Baseline infection</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>844</td>
<td>3.45</td>
</tr>
<tr>
<td>Yes</td>
<td>449</td>
<td>3.78</td>
</tr>
<tr>
<td>Secondary AML</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>978</td>
<td>3.46</td>
</tr>
<tr>
<td>Yes</td>
<td>312</td>
<td>3.91</td>
</tr>
<tr>
<td>Prior chemotherapy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1,098</td>
<td>3.53</td>
</tr>
<tr>
<td>Yes</td>
<td>195</td>
<td>3.76</td>
</tr>
</tbody>
</table>

### Continuous variables

<table>
<thead>
<tr>
<th>Continuous variables</th>
<th>Spearman correlation (any age)$^*$</th>
<th>$P$</th>
<th>Spearman correlation (age $\geq 60$ y)$^*$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.30</td>
<td>&lt;0.0001</td>
<td>0.16</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Zubrod performance status</td>
<td>0.29</td>
<td>&lt;0.0001</td>
<td>0.27</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>WBC</td>
<td>0.33</td>
<td>&lt;0.0001</td>
<td>0.44</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Creatinine</td>
<td>0.52</td>
<td>&lt;0.0001</td>
<td>0.54</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Creatinine clearance</td>
<td>-0.46</td>
<td>&lt;0.0001</td>
<td>-0.45</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Albumin</td>
<td>-0.29</td>
<td>&lt;0.0001</td>
<td>-0.25</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>LDH</td>
<td>0.29</td>
<td>&lt;0.0001</td>
<td>0.37</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Alkaline phosphatase</td>
<td>0.21</td>
<td>&lt;0.0001</td>
<td>0.19</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Uric acid</td>
<td>0.39</td>
<td>&lt;0.0001</td>
<td>0.40</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>ANC</td>
<td>0.25</td>
<td>&lt;0.0001</td>
<td>0.34</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Prothrombin time</td>
<td>0.35</td>
<td>&lt;0.0001</td>
<td>0.37</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Circulating blasts</td>
<td>0.15</td>
<td>&lt;0.0001</td>
<td>0.23</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Circulating monocytes</td>
<td>0.25</td>
<td>&lt;0.0001</td>
<td>0.23</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

**NOTE:** Associations of $\beta_2$M levels with other pretreatment characteristics were assessed by t test if two categories were present and by ANOVA test if more than two categories were present for each variable.

*Spearman rank correlation coefficient assesses associations between $\beta_2$M levels and other continuous variables.
Diagnosis. The diagnosis of AML was confirmed at M. D. Anderson by an M. D. Anderson hematopathologist after review of bone marrow aspiration and/or peripheral blood smear. Conventional cytogenetic studies were done on bone marrow aspirate material using standard G-banding techniques. FLT3 analysis was done using a fluorescent multiplex PCR and restriction digestion method followed by capillary electrophoresis to detect internal tandem duplications and D835Y point mutations in the FLT3 gene. FLT3 was positive if either internal tandem duplication or D835Y mutation was identified.

Endpoints and statistical methods. We were primarily interested in the association between β2M levels and survival time. To feel more confident that any relationship between β2M and survival did not reflect differences in therapy given after relapse between patients with higher and lower β2M levels, we also analyzed event-free survival (EFS; i.e., survival from the start of treatment until relapse or death). Complete remission (CR) was defined using standard criteria, for example, a morphologically leukemia-free state, including evidence of normal erythropoiesis, granulopoiesis, and megakaryocytopoiesis; and ≤5% blasts in bone marrow aspirate, an absolute neutrophil count >1 × 10⁹/L, and platelets ≥100 × 10⁹/L (38). Relapse-free survival (RFS) was measured from the CR date until relapse or death.

Initially, we assessed the association between β2M levels and CR, overall survival, EFS, and RFS by dividing the patients into four groups with β2M levels of <1, 1-1.5, 1.5-2.0, and >2.0 times the upper limit of normal. The relationship between β2M levels and other covariates was examined using the t test, ANOVA, or the Spearman rank correlation coefficient, as appropriate. Overall survival, EFS, and RFS and logistic regression for CR and included the following covariates: age, sex, race; performance status; cytogenetics; WBC count; hemoglobin level; platelet count; absolute neutrophil count; proportion of circulating monocytes, neutrophils, blasts, metamyelocytes, and myelocytes; creatinine clearance; levels of serum creatinine, bilirubin, lactate dehydrogenase (LDH), β2M, albumin, uric acid, glucose, alkaline phosphatase, alanine aminotransferase, and fibrinogen; prothrombin time and partial prothrombin time; FLT-3 and RAS mutations; French-American-British subtype; proportion of bone marrow blasts; proportion of CD2, CD33, and CD56 in bone marrow blasts by immunophenotyping; secondary versus de novo AML; history of antecedent hematologic disorder (history of a hemoglobin level <12 g/dL), a platelet count <150,000/μL, a neutrophil count <1,500/μL, or a WBC count >20,000/μL, or for at least 1 month before M. D. Anderson presentation) or other malignancy; presence of infection; ara-C–containing therapy; and days from diagnosis to treatment. A stepwise variable selection procedure was done to identify independent variables, which determined the final model. Patients with missing data were excluded from the multivariate analyses; in these analyses, numerical covariates (age, β2M, hemoglobin, WBC, etc.) were considered as such.

We analyzed patients ≥60 years separately from those younger than 60 years. Given the large number of tests of significance done and the risk of false-positive results, we divided our patients into two independent groups, using the first (training) to derive and the second (validation) to test the model. We also used a “bootstrapping” method in which we repeatedly (100 times) randomly selected 50% of the total population and assessed in which proportion of the 100 samples the Cox model identified a covariate as having independent prognostic significance. Those covariates found to be significant in the training and validation sets and/or noted to be independently significant at P < 0.05 in ≥40% of the randomly selected 100 bootstrapping sets were used to derive a scoring algorithm that can be used to predict a given patient’s risk of death. In general, the scores assigned the different covariates were determined using the estimated coefficients from the fitted Cox models.
model. Because such a scoring system assumes arbitrarily that there are cutoffs for numerical variables, we also used a method to validate the scoring system in which a continuous range of scores is generated (39). Statistical analyses were carried out using SAS 8.2 and SPLUS 2000 (Insightful Corporation).

Results

Correlation of β2M levels with other covariates. High β2M levels were correlated with male gender (P < 0.0001), intermediate and poor-risk cytogenetics (P = 0.03), presence of RAS mutations (P = 0.003), baseline infection (P = 0.04), and secondary AML (P = 0.01; Table 1). High β2M levels also were associated with older age (P < 0.0001) and other factors. Table 1 also illustrates these associations in patients 60 years or older.

Patient characteristics of patients with untreated AML with and without measurements of serum β2M levels are shown in Table 2. After adjusting for covariates (cytogenetics, age, Zubrod performance status, secondary versus de novo AML, history of antecedent hematologic disorder, levels of hemoglobin, prothrombin time, LDH, uric acid, and creatinine and ara-C–containing therapy versus therapy without ara-C), patients in whom β2M levels were measured had longer survival, reflecting that failure to measure β2M levels was

Table 3. Response, overall survival, and RFS by pretreatment characteristics

<table>
<thead>
<tr>
<th>n</th>
<th>CR, %</th>
<th>P</th>
<th>Median survival, mo</th>
<th>P</th>
<th>Median RFS, mo</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y</td>
<td>&lt;60</td>
<td>590</td>
<td>69</td>
<td>&lt;0.0001</td>
<td>17</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>&gt;60</td>
<td>703</td>
<td>47</td>
<td>6</td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cytogenetics, risk</td>
<td>Favorable</td>
<td>87</td>
<td>91</td>
<td>&lt;0.0001</td>
<td>56</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Intermediate</td>
<td>781</td>
<td>60</td>
<td>13</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poor</td>
<td>374</td>
<td>42</td>
<td>5</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>β2M (mg/L)</td>
<td>&lt;2.0</td>
<td>251</td>
<td>67</td>
<td>16</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>2.0-2.9</td>
<td>448</td>
<td>66</td>
<td>14</td>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.0-3.9</td>
<td>260</td>
<td>54</td>
<td>8</td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥4.0</td>
<td>334</td>
<td>39</td>
<td>5</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PS</td>
<td>0-1</td>
<td>947</td>
<td>61</td>
<td>13</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>&gt;1</td>
<td>346</td>
<td>44</td>
<td>4</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LDH</td>
<td>&lt;1.5 × ULN</td>
<td>661</td>
<td>60</td>
<td>12</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>≥1.5 × ULN</td>
<td>624</td>
<td>54</td>
<td>8</td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prior cancer</td>
<td>Yes</td>
<td>312</td>
<td>46</td>
<td>8</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>978</td>
<td>60</td>
<td>11</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Creatinine, mg/dL</td>
<td>≤1.6</td>
<td>1191</td>
<td>59</td>
<td>11</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>&gt;1.6</td>
<td>101</td>
<td>30</td>
<td>2</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Creatinine clearance, mL/min*</td>
<td>≥70</td>
<td>647</td>
<td>51</td>
<td>7.2</td>
<td>10.4</td>
<td></td>
</tr>
<tr>
<td>&lt;70</td>
<td>646</td>
<td>63</td>
<td>13.2</td>
<td>11.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leukocytes, ×10⁹/L</td>
<td>&lt;10</td>
<td>709</td>
<td>60</td>
<td>12</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>≥10</td>
<td>584</td>
<td>53</td>
<td>8</td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Albumin, g/dL</td>
<td>&lt;3.5</td>
<td>713</td>
<td>53</td>
<td>7</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>≥3.5</td>
<td>578</td>
<td>62</td>
<td>14</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PT, s</td>
<td>≤Median</td>
<td>613</td>
<td>64</td>
<td>14</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>&gt;Median</td>
<td>661</td>
<td>51</td>
<td>7</td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bilirubin, g/dL</td>
<td>&lt;1</td>
<td>1059</td>
<td>58</td>
<td>10</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>≥1</td>
<td>233</td>
<td>49</td>
<td>7</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alkaline phosphatase, IU/L</td>
<td>&lt;126</td>
<td>1079</td>
<td>58</td>
<td>11</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>≥126</td>
<td>208</td>
<td>52</td>
<td>6</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FAB subtype</td>
<td>AMOL+AMML</td>
<td>355</td>
<td>62</td>
<td>9</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>938</td>
<td>55</td>
<td>10</td>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FLT-3 mutations</td>
<td>Yes</td>
<td>134</td>
<td>58</td>
<td>7</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>428</td>
<td>56</td>
<td>14</td>
<td>12</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: PS, performance status; ULN, upper limit of normal; PT, prothrombin time; FAB, French-American-British; AMOL, acute monocytic leukemia; AMML, acute myelomonocytic leukemia.

*Creatinine clearance was calculated using the Cockcroft and Gault equation.
Fig. 1. Survival in patients with AML by (A) $\beta_2 M$ levels ($n = 1,293$); (B) $\beta_2 M$ levels in patients 60 y or older ($n = 703$); (C) age; (D) cytogenetic risk; favorable, intermediate, poor, or unknown; (E) performance status; and (F) Kaplan-Meier survival curves by prognostic score in patients ≥60 y. The risk of death could be determined by summing the scores present at diagnosis, and ranged from 0 to 8. Cytogenetics (0, 1, or 2); performance status (0 or 1); and levels of $\beta_2 M$ (0 or 1), uric acid (0 or 1), and LDH (0 or 1), formed the basis of a scoring system (Table 5A). At 1 y, 63%, 46%, 41%, 30%, 17%, and 4% of patients with scores 0, 1, 2, 3, 4, and 5 to 6, respectively, are expected to be alive. G, Kaplan-Meier survival curves by prognostic score in patients ≥60 y. The risk of death could be determined by summing the scores present at diagnosis, and ranged from 0 to 8. In patients younger than 60 y, cytogenetics (0, 2, or 4); performance status (0 or 1); age (0 or 1), hemoglobin level (0 or 1), and leukocyte count (0 or 1) were used to determine a patient’s score (Table 6). At 1 y, 95%, 76%, 71%, 66%, 39%, and 24% of patients with scores 0-1, 2, 3, 4, 5, and 6-8, respectively, are expected to be alive.
frequently due to the need to treat a sick patient on an emergency basis.

**Patient characteristics.** We focused on the 1,293 patients whose β2M levels were measured. The median age of these patients was 61 years (range, 16-89 years). Zubrod performance status was 0 in 13.8%, 1 in 59.5%, 2 in 19.7%, 3 in 5%, and 4 in 2%. The distribution of French-American-British subtypes was as follows: M0, 7.5%; M1, 17%; M2, 33%; M4, 22%; M5, 10%; M6, 6%; M7, 2.5%; and other, 2%. The median β2M level was 2.8 mg/L (range, 0.7-31.3 mg/L). Overall, 541 of the 1,293 (42%) patients had a history of antecedent hematologic disorder. In patients with such a history, the mean duration of antecedent hematologic disorder was 0.75 years (SD, 2.6 years). Of the patients, 96% received cytotoxic chemotherapy, which contained ara-C in 91% of these cases, and 4% of patients received “targeted therapy.” The CR rate was 57%. The median follow-up of surviving patients was 27 months (range, 1-145 months). The median overall survival, RFS, and EFS durations were 9.9, 12.2, and 4.2 months, respectively.

**Response.** Table 3 shows CR rates according to pretreatment characteristics, including β2M levels. In multivariate analysis, independent factors predicting response were younger age (P < 0.0001), better-risk cytogenetics (P < 0.0001), lower β2M levels (P < 0.0001), lower creatinine clearance levels (P = 0.0001), de novo (versus secondary) AML (P = 0.01), and better Zubrod performance status (P = 0.02).

**Overall survival.** When categorized in relation to the upper limit of normal, as described above, lower serum β2M levels predicted longer survival (Fig. 1A). Figure 1B shows survival by serum β2M levels in patients 60 years or older. Survival by age, cytogenetic risk group, and performance status is shown in Fig. 1C to E. Among patients younger than 60 years, with elevated B2 microglobulin (>2 mg/dL), 15 of 299 (5%) underwent stem cell transplantation as maintenance therapy in first CR. The median follow-up is 2.94 years for transplanted patients and 3.27 years for nontransplanted. Six of the 15 patients who underwent stem cell transplantation and 158 of the 284 who did not receive stem cell transplantation have died. The median survival of the 15 patients was 2.58 years compared with 1.96 years in the 284 patients who did not undergo stem cell transplantation (P = 0.37).

**Table 4. Multivariate analysis using Cox model showing interaction between age and β2M in 1,293 patients**

<table>
<thead>
<tr>
<th>Variable</th>
<th>RR (95% CI)</th>
<th>Z score</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Worsening-risk CG</td>
<td>2.06 (1.82-2.34)</td>
<td>11.33</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Performance status</td>
<td>1.24 (1.14-1.35)</td>
<td>4.86</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Secondary vs de novo AML</td>
<td>1.50 (1.26-1.79)</td>
<td>4.54</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Uric acid</td>
<td>1.03 (1.01-1.04)</td>
<td>4.10</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Prothrombin time</td>
<td>1.08 (1.06-1.13)</td>
<td>3.59</td>
<td>0.0003</td>
</tr>
<tr>
<td>LDH</td>
<td>1.06 (1.03-1.10)</td>
<td>3.57</td>
<td>0.0004</td>
</tr>
<tr>
<td>Hemoglobin</td>
<td>0.94 (0.91-0.98)</td>
<td>-3.06</td>
<td>0.002</td>
</tr>
<tr>
<td>Ara-C-containing therapy</td>
<td>0.77 (0.63-0.94)</td>
<td>-2.56</td>
<td>0.01</td>
</tr>
<tr>
<td>Presence of infection</td>
<td>1.20 (1.04-1.39)</td>
<td>2.45</td>
<td>0.01</td>
</tr>
<tr>
<td>Age</td>
<td>1.02 (1.01-1.02)</td>
<td>4.02</td>
<td>0.0001</td>
</tr>
<tr>
<td>β2M</td>
<td>0.93 (0.82-1.04)</td>
<td>-1.27</td>
<td>0.20</td>
</tr>
<tr>
<td>Age: β2M</td>
<td>1.00 (1.00-1.00)</td>
<td>2.03</td>
<td>0.04</td>
</tr>
</tbody>
</table>

**Event-free survival.** Serum β2M levels also were associated with the duration of EFS (n = 1,293). In multivariate analysis, independent factors predicting longer EFS were better-risk cytogenetics (P < 0.0001), younger age (P < 0.0001), lower β2M levels (P = 0.0002), de novo (versus secondary) AML (P = 0.0002), lower uric acid levels (P = 0.0004), higher hemoglobin levels (P = 0.002), lower LDH levels (P = 0.003), better performance status (P = 0.007), shorter prothrombin time (P = 0.01), and absence of baseline infection (P = 0.04).

**Relapse-free survival.** Serum levels of β2M were also associated with duration of RFS (Table 3). Patients with β2M levels >4 mg/L had inferior rates of RFS, although the RFS curves of the other β2M groups started to overlap after 1 to 2 years. Independent factors predicting longer RFS were better-risk cytogenetics (P < 0.0001), younger age (P < 0.0001), higher hemoglobin levels (P = 0.01), lower β2M levels (P = 0.001), and de novo (versus secondary) AML (P = 0.01).

**Prognostic interaction between β2M levels and age.** The associations between β2M level and many of the factors that were also related to survival prompted us to perform a multivariate analysis and fit a multivariate model, which included the interaction between β2M and age. According to this analysis, the prognostic significance of β2M increased with age, as evidenced by the P value of 0.04 for the interaction between age and β2M (Table 4). The P value of 0.20 for β2M reflects the fact that β2M was not prognostically significant in younger patients. The interaction between β2M and age led us to develop separate models for patients <60 and ≥60 years. Because ara-C–treated patients had longer survival (Table 4), each of these models was restricted to ara-C–treated patients.

**Patients ≥60 years.** The 591 ara-C–treated patients ≥60 years had a median age of 70 years. Fifty percent of the patients (n = 296) were used to develop a model and the remaining 50% (n = 295) to validate the model. Cytogenetics, performance status, and β2M and uric acid levels were each an independent factor in both the training and validation sets (Table 5A), whereas LDH was prognostic only in the population used to develop, not test, the model. Next, we resampled 100 times the 591 patients to obtain separate 296-patient data sets, in each of which we fit a Cox model. Cytogenetics and performance status were independently significant in 100% and 98% of these sets. Corresponding figures for β2M, LDH, and uric acid were 48%, 45%, and 42%. In contrast, age above and below the median of 70 years was independently significant in only 15 of the 100 samples; corresponding values for secondary (versus de novo) AML, hemoglobin, bilirubin, WBC, infection, albumin, and creatinine were 39%, 25%, 13%, 10%, 5%, 4%, and 4%, respectively.

To determine a scoring system, we included those covariates that were predictive in >40% of the 100 repeated samples. Four of these five covariates (performance status, cytogenetics, β2M, and uric acid) were also predictive in the 295-patient test set (Table 5A). Having selected our covariates, we included all 591 patients age ≥60 years (Table 5A) to develop the scoring system, using the relative risk from the Cox model to define the score, which could range from 0 to 6.

Application of this scoring system is shown in Table 5A and Fig. 1F. We compared the scoring system with a more formal system defined without using a cutoff for numerical values.
such as $\beta_2$M, as previously described (39). The more informal and the formal methods gave similar results. Because the results of cytogenetic analysis are not often available when a treatment decision must be made, we also did a multivariate analysis for survival without including cytogenetics. Independent factors predicting longer survival were performance status 0 to 1 ($P < 0.0001$), de novo (versus secondary) AML ($P = 0.006$), $\beta_2$M $< 3$ mg/dL ($P = 0.0097$), uric acid less than the upper limit of normal ($P = 0.012$), LDH $< 1.5 \times ULN$ ($P = 0.014$), and age $< 70$ years ($P = 0.025$; Table 5B).

Patients $< 60$ years. Among 589 ara-C–treated patients $< 60$ years, the median age was 48 years. Fifty percent of the patients ($n = 295$) were used to develop a training set and the remaining 50% ($n = 294$) to validate the model. Cytogenetics, performance status, older age, and higher WBC counts each independently affected survival in both the training and validation sets (Table 6), whereas lower hemoglobin level was prognostic in the training but not in the validation set. When the 589 patients were resampled to obtain 100 separate 295-patient data sets, in each of which we fit a Cox model, cytogenetics were independently significant in 99% of the samples, age in 83%, hemoglobin in 53%, WBC in 52%, and performance status $> 1$ in 45%. $\beta_2$M levels were significant in only 8% of the samples. The respective proportions for LDH, infection, creatinine, prothrombin time, and history of cancer were 39%, 22%, 4%, 3%, and 1%. On the

<table>
<thead>
<tr>
<th>Table 5. Factors independently prognostic of survival and scoring system in ara-G treated patients $\geq 60$ y</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. Factors independently prognostic of survival</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Training sample ($n = 296$)</strong></td>
</tr>
<tr>
<td>Worse-risk cytogenetics</td>
</tr>
<tr>
<td>Poorer performance status</td>
</tr>
<tr>
<td>$\beta_2$M</td>
</tr>
<tr>
<td>LDH</td>
</tr>
<tr>
<td>Uric acid</td>
</tr>
<tr>
<td>Validation sample ($n = 295$)</td>
</tr>
<tr>
<td>Worse-risk cytogenetics</td>
</tr>
<tr>
<td>Poorer performance status</td>
</tr>
<tr>
<td>$\beta_2$M</td>
</tr>
<tr>
<td>LDH</td>
</tr>
<tr>
<td>Uric acid</td>
</tr>
<tr>
<td>RR (95% CI)</td>
</tr>
<tr>
<td>Z score</td>
</tr>
<tr>
<td>$P$</td>
</tr>
<tr>
<td>Scoring system for survival model</td>
</tr>
<tr>
<td>Covariate</td>
</tr>
<tr>
<td>Relative risk</td>
</tr>
<tr>
<td>Score</td>
</tr>
<tr>
<td>Cytogenetic, risk</td>
</tr>
<tr>
<td>Poor</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>Performance status</td>
</tr>
<tr>
<td>$&gt; 1$</td>
</tr>
<tr>
<td>0-1</td>
</tr>
<tr>
<td>$\beta_2$M, mg/dL</td>
</tr>
<tr>
<td>$&gt; 3$</td>
</tr>
<tr>
<td>$&lt; 3$</td>
</tr>
<tr>
<td>Uric acid</td>
</tr>
<tr>
<td>$&gt; ULN$</td>
</tr>
<tr>
<td>$&lt; ULN$</td>
</tr>
<tr>
<td>LDH</td>
</tr>
<tr>
<td>$&gt; 1.5 \times ULN$</td>
</tr>
<tr>
<td>$&lt; 1.5 \times ULN$</td>
</tr>
<tr>
<td><strong>B. Factors (without cytogenetics) independently prognostic of survival in 591 patients $\geq 60$ y with newly diagnosed AML</strong></td>
</tr>
<tr>
<td>RR (95% CI)</td>
</tr>
<tr>
<td>Z score</td>
</tr>
<tr>
<td>$P$</td>
</tr>
<tr>
<td>PS 2 vs 0-1</td>
</tr>
<tr>
<td>PS 3-4 vs 0-1</td>
</tr>
<tr>
<td>Secondary vs de novo AML</td>
</tr>
<tr>
<td>$\beta_2$M $&gt; 3$ mg/dL</td>
</tr>
<tr>
<td>Uric acid $&gt; ULN$</td>
</tr>
<tr>
<td>LDH $&gt; 1.5 \times ULN$</td>
</tr>
<tr>
<td>Age $&gt; 70$ y</td>
</tr>
</tbody>
</table>
basis of the results of the resamples, and the relative risk from the Cox model, which included all 589 patients (Table 6), poor-risk cytogenetics, age ≥48 years, performance status >1, WBC counts ≥10 x 10^9/L, and hemoglobin <8 g/dL were used to develop a scoring system for patients <60 years old (Table 6). Patient scores can range from 0 to 8. Application of this scoring system is shown in Fig. 1G.

### Discussion

This is the first study to explore whether serum β2M levels are an independent risk factor in AML, as is the case in various other hematologic malignancies. Our data strongly support the independent prognostic role of uncorrected pretreatment serum β2M levels in patients with AML who are ≥60 years old and are given ara-C–containing therapy. Notably, β2M levels did not predict survival in patients younger than 60 years. This observation is consistent with qualitative differences in the AML of older and younger patients.

In multivariate analysis using training, validation, and resampling sets, poor-risk cytogenetics, performance status >1, β2M >3 mg/dL, uric acid greater than the upper limit of normal, and LDH >1.5 x the upper limit of normal were the top five adverse independent factors predicting survival in patients >60 years. Similarly, poor-risk cytogenetics, performance status >1, age ≥48 years, hemoglobin <8 g/dL, and leukocyte count ≥10 x 10^9/L were the top five adverse independent factors predicting survival in patients <60 years. To optimize the clinical use of these findings, two simple prognostic scores were developed, one for patients ≥60 years and another for patients <60 years.

A fundamental question is whether our results can be generalized to other patients with AML. The other covariates that predicted outcome in our patients were, by and large, those noted in many other series, and the outcomes themselves (median survival, long-term survival, etc.) were similar to those often reported elsewhere. On the other hand, it is important to bear in mind that β2M measurement was an independent prognostic factor; that is, our patients in whom β2M was not measured did worse than patients in whom β2M was measured. This observation presumably reflects the “sicker” nature of the former patients and suggests caution in the application of our system to such patients.

### Table 6. Factors independently prognostic of survival and scoring system in ara-C–treated patients <60 y

<table>
<thead>
<tr>
<th>Covariate</th>
<th>RR (95% CI)</th>
<th>Z score</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Training sample (n = 295)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Worse-risk cytogenetics</td>
<td>2.12 (1.63-2.76)</td>
<td>5.60</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Hemoglobin</td>
<td>0.88 (0.81-0.96)</td>
<td>-3.08</td>
<td>0.002</td>
</tr>
<tr>
<td>Performance status &gt;1</td>
<td>1.63 (1.13-2.33)</td>
<td>2.64</td>
<td>0.008</td>
</tr>
<tr>
<td>Age</td>
<td>1.02 (1.00-1.03)</td>
<td>2.30</td>
<td>0.02</td>
</tr>
<tr>
<td>WBC</td>
<td>1.00 (1.00-1.006)</td>
<td>2.30</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Validation sample (n = 294)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Worse-risk cytogenetics</td>
<td>2.70 (2.03-3.59)</td>
<td>6.81</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>WBC</td>
<td>1.00 (1.00-1.01)</td>
<td>2.99</td>
<td>0.003</td>
</tr>
<tr>
<td>Age</td>
<td>1.03 (1.00-1.03)</td>
<td>2.10</td>
<td>0.04</td>
</tr>
<tr>
<td>Performance status &gt;1</td>
<td>1.58 (1.12-2.23)</td>
<td>2.58</td>
<td>0.009</td>
</tr>
<tr>
<td>Hemoglobin</td>
<td>0.97 (0.90-1.04)</td>
<td>-0.87</td>
<td>0.38</td>
</tr>
<tr>
<td>All patients (n = 589)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poor- vs favorable-risk cytogenetics</td>
<td>5.44 (3.49-8.45)</td>
<td>7.51</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Intermediate- vs favorable-risk cytogenetics</td>
<td>2.23 (1.48-3.37)</td>
<td>3.83</td>
<td>0.0001</td>
</tr>
<tr>
<td>Age ≥48 y</td>
<td>1.38 (1.12-1.71)</td>
<td>2.97</td>
<td>0.003</td>
</tr>
<tr>
<td>PS 2 vs 0-1</td>
<td>1.71 (1.31-2.23)</td>
<td>3.92</td>
<td>0.0001</td>
</tr>
<tr>
<td>PS 3-4 vs 0-1</td>
<td>1.54 (1.02-2.32)</td>
<td>2.05</td>
<td>0.04</td>
</tr>
<tr>
<td>Hemoglobin &lt;8 g/dL</td>
<td>1.31 (1.05-1.62)</td>
<td>2.45</td>
<td>0.01</td>
</tr>
</tbody>
</table>

### Scoring system for survival model

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Relative risk</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cytogenetic, risk</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poor</td>
<td>5.44</td>
<td>4</td>
</tr>
<tr>
<td>Intermediate</td>
<td>2.23</td>
<td>2</td>
</tr>
<tr>
<td>Favorable</td>
<td>1.00</td>
<td>0</td>
</tr>
<tr>
<td>Performance status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;1</td>
<td>1.71</td>
<td>1</td>
</tr>
<tr>
<td>0-1</td>
<td>1.00</td>
<td>0</td>
</tr>
<tr>
<td>Age, y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥48</td>
<td>1.38</td>
<td>1</td>
</tr>
<tr>
<td>&lt;48</td>
<td>1.00</td>
<td>0</td>
</tr>
<tr>
<td>Hemoglobin, g/dL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;8</td>
<td>1.31</td>
<td>1</td>
</tr>
<tr>
<td>≥8</td>
<td>1.00</td>
<td>0</td>
</tr>
<tr>
<td>WBC, x 10^9/L</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥10</td>
<td>1.28</td>
<td>1</td>
</tr>
<tr>
<td>&lt;10</td>
<td>1.00</td>
<td>0</td>
</tr>
</tbody>
</table>

Imaging, Diagnosis, Prognosis
Another potential problem with generalizing our findings relates to the many tests of statistical significance we did and the consequent possibility of false-positive results. Indeed, whereas the use of multivariate models is common, they are often not validated. The prognostic models presented here were validated in two ways: training and validation sets (50% each) and multiple resampling (“bootstrapping”). Although it could be argued that the decision to select only those factors that were found to be independently prognostic at P < 0.05 in at least 40% of the resamples is arbitrary, we would point out that in general these factors were also found to be significant in the validation set, which used two totally independent populations.

Another observation that emerged from the current study is the association between elevated serum levels of β2M and other covariates. These observations are in line with previous reports showing that β2M levels are elevated in normal individuals who are older, male, and have elevated creatinine levels. Although no association between β2M levels and monocytic AML by the French-American-British classification was noted, as previously described (13, 36, 37, 40–42), we found that elevated β2M levels were correlated with increased numbers of monocytes. Several of the correlations, such as those between β2M levels and leukocyte counts, circulating blasts or monocytes, and LDH indicate that β2M levels are a marker of increased turnover. These correlations were also significant in patients 60 years or older.

The clinical significance of serum β2M levels in AML may be explained as previously described in non–Hodgkin’s lymphoma (14). β2M and HLA molecules are variably expressed on the surface of tumor cells (43–46). Poor expression of MHC class I molecules is considered an adverse prognostic factor in lymphomas (47–49). Because the HLA class I antigens are key components of the immune system, the recognition of tumor-specific antigens by cytotoxic T cells is dependent on intact HLA expression. Malignant cells with modified or lacking HLA expression may escape from the normal host immune response and proliferate uncontrollably (50–52). A structural defect of the HLA complex may also result in changes in epitope expression and increased release of β2M in serum (8, 9, 53, 54). This mechanism is consistent with the observation that serum β2M levels reflect tumor burden and cell turnover (7–10).

The prognostic significance of serum β2M levels has particular interest in light of recent studies in mouse models that showed that monoclonal antibodies specific to human β2M induced apoptosis in vitro and had antitumor activity against multiple myeloma, Burkitt’s lymphoma, mantle cell lymphoma, T-cell leukemia, and myeloid leukemia (55). These effects had selective antitumor activity without damaging marrow hematopoietic cells of implanted human bone or marrow organs that express human β2M/HLA-A2 molecules, providing a potential therapeutic approach for patients with AML or other hematologic malignancies with elevated serum β2M levels (55). The absence of damage to normal hematopoietic cells may be explained by the activity of the monoclonal antibody against cells with structural defects of the HLA complex, such as tumor cells, but not against normal cells with intact HLA structure.

In conclusion, our data show that serum β2M levels are highly predictive for clinical outcomes in patients ≥60 years with newly diagnosed AML. Given this observation and the relative ease with which β2M can be assessed, the proposed scoring system seems to represent a significant advantage in staging these patients to recognize those who may benefit from different treatment approaches and to allow a more accurate comparison between different therapeutic approaches.

References

The Prognostic Significance of Serum $\beta_2$ Microglobulin Levels in Acute Myeloid Leukemia and Prognostic Scores Predicting Survival: Analysis of 1,180 Patients

Apostolia-Maria Tsimberidou, Hagop M. Kantarjian, Sijin Wen, et al.


Updated version
Access the most recent version of this article at:
http://clincancerres.aacrjournals.org/content/14/3/721

Cited articles
This article cites 55 articles, 20 of which you can access for free at:
http://clincancerres.aacrjournals.org/content/14/3/721.full#ref-list-1

Citing articles
This article has been cited by 4 HighWire-hosted articles. Access the articles at:
http://clincancerres.aacrjournals.org/content/14/3/721.full#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.