Association of a Leukemic Stem Cell Gene Expression Signature With Clinical Outcomes in Acute Myeloid Leukemia

Andrew J. Gentles, PhD
Sylvia K. Plevritis, PhD
Ravindra Majeti, MD, PhD
Ash A. Alizadeh, MD, PhD

Acute myeloid leukemia (AML) is an aggressive malignancy of the bone marrow characterized by accumulation of early myeloid blood cells that fail to mature and differentiate. The course of the disease is marked by poor prognosis, frequent relapse, and high disease–related mortality. Recent clinical investigation has focused on the identification of prognostic subgroups in adult AML with the goal of guiding patients into risk-adapted therapies. Such investigation determined that cytogenetic abnormalities are prognostic, some favorable and others unfavorable, yet up to 50% of patients have normal karyotype AML with a wide range of clinical outcomes. In these patients, the presence of specific molecular mutations can provide prognostic information, including internal tandem duplications within the FLT3 gene, partial tandem duplication of the MLL gene, mislocalizing mutations of the NPM1 gene, mutations in the CEBPA and RAS genes, and increased expression of the BAALC and ERG genes. However, these parameters and others such as patient age are only partially successful at capturing risk of relapse and patient outcomes following treatment.

A growing body of evidence suggests that specific cancer cell subpopulations in AML have unique capabilities of initiating and maintaining tumors. The strongest support for this cancer stem cell model comes from transplantation assays in immunodeficient mice, which indicate that human acute myeloid leukemia (AML) is driven by self-renewing leukemic stem cells (LSCs). This model has significant implications for the development of novel therapies, but its clinical relevance has yet to be determined.

Objective To identify an LSC gene expression signature and test its association with clinical outcomes in AML.

Design, Setting, and Patients Retrospective study of global gene expression (microarray) profiles of LSC-enriched subpopulations from primary AML and normal patient samples, which were obtained at a US medical center between April 2005 and July 2007, and validation data sets of global transcriptional profiles of AML tumors from 4 independent cohorts (n=1047).

Main Outcome Measures Identification of genes discriminating LSC-enriched populations from other subpopulations in AML tumors; and association of LSC-specific genes with overall, event-free, and relapse-free survival and with therapeutic response.

Results Expression levels of 52 genes distinguished LSC-enriched populations from other subpopulations in cell-sorted AML samples. An LSC score summarizing expression of these genes in bulk primary AML tumor samples was associated with overall survival as a continuous variable (hazard ratio [HR], 1.15; 95% confidence interval [CI], 1.08-1.22; log-likelihood P<.001). The absolute risk of death by 3 years was 57% (95% CI, 43%-67%) for the low LSC score group compared with 78% (95% CI, 66%-86%) for the high LSC score group (HR, 1.9 [95% CI, 1.3-2.7]; log-rank P=.002). In another cohort with available data on event-free survival for 70 patients with normal karyotypes, the risk of an event by 3 years was 48% (95% CI, 27%-63%) in the low LSC score group vs 81% (95% CI, 60%-91%) in the high LSC score group (HR, 2.4 [95% CI, 1.3-4.5]; log-rank P=.006). In multivariate Cox regression including age, mutations in FLT3 and NPM1, and cytogenetic abnormalities, the HRs for LSC score in the 3 cohorts with data on all variables were 1.07 (95% CI, 1.01-1.13; P=.02), 1.10 (95% CI, 1.03-1.17; P=.005), and 1.17 (95% CI, 1.05-1.30; P=.005).

Conclusion High expression of an LSC gene signature is independently associated with adverse outcomes in patients with AML.

See also Patient Page.
cell sorting using the antibodies shown in eFigure 1 as follows: AML LSCs (n=7), AML LPCs (n=7), AML blasts (n=7), normal hematopoietic stem cells (HSCs) (bone marrow and cord blood, n=7), normal multipotent progenitors (bone marrow and cord blood, n=8), normal common myeloid progenitors (bone marrow, n=4), normal granulocyte-monocyte progenitors (bone marrow, n=4), and megakaryocyte-erythroid progenitors (bone marrow, n=4). Global transcriptional profiles were generated for each sample using Affymetrix U133 Plus 2.0 gene expression microarrays (Affymetrix, Santa Clara, California). Raw data were deposited at the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus GSE24006. The detailed experimental procedures for the purification of cell subpopulations have been reported previously. Global gene expression profiles of 14 paired LSC-enriched and LPC-enriched subpopulations purified from the 7 AML patient samples from Stanford University Medical Center were combined with 16 paired profiles (8 LSC-enriched and 8 LPC-enriched subpopulations) from an independent study to produce 1 data set of 30 samples for these analyses. All genes profiled on microarrays were ranked by the mean ratio of their expression between paired LSC-enriched and LPC-enriched subpopulations, and evaluated using Gene Set Enrichment Analysis. This approach assessed whether any predefined groups of biologically related genes were concordantly more highly expressed in LSC-enriched or LPC-enriched subpopulations. Individual genes expressed more highly in LSC-enriched subpopulations compared with LPC-enriched subpopulations (or vice versa) were identified using Significance Analysis of Microarrays (false discovery rate <10%). Ingenuity Pathways Analysis was used to identify interaction networks involving these genes. The genes that were more highly expressed in the LSC-enriched subpopulations relative to the LPC-enriched subpopulations were summarized by an LSC signature. The LSC signature was defined as a single number representing the relative expression of the LSC-enriched genes in a given sample compared with the other samples in the same data set. The signature was computed as the first principal component of the gene expression data matrix whose rows were the LSC-enriched genes. Each column of the matrix represented 1 sample, and matrix entries were the gene expression values corresponding to each gene in each sample. By definition, the first principal component of such a data matrix is the weighted sum of the genes’ expression levels that explains the maximum possible amount of their total variation across all samples. Thus, for a group of n genes, each sample’s weight was then applied to the test cohort. The LSC signature was evaluated in the purified normal and leukemic subpopulations to investigate the expression of the LSC-enriched genes beyond the LSC-enriched and LPC-enriched subpopulations used to identify them.

To test associations between the LSC-enriched genes and clinical outcomes, a retrospective training–validation scheme was adopted. Raw microarray data were obtained for 4 publicly available bulk AML gene expression studies with available clinical annotations from the NCBI Gene Expression Omnibus (GSE12417 (n=163 normal karyotype AML only with overall survival outcomes), GSE10358 (n=184 with overall survival and event-free survival known for 178 cases), GSE14468 (n=526 with overall survival and event-free survival known for 262 cases and relapse-free survival known for 213 cases) and from the National Cancer Institute caArray database (willm-00119 [https://array.nci.nih.gov/caarray/project/willm-00119], n=170 with overall survival only). The largest normal karyotype AML data set (n=163) was used as the training set and the other 3 data sets were used for validation. The LSC signature was calculated in the training cohort and the same weights were then applied to the test cohorts. Because the gene weights were not
recomputed in each test cohort to ensure unbiased validation, the resulting number for each sample was referred to as the LSC score. In the training set, the LSC signature and LSC score were identical by definition. The median LSC score in the training set was used to partition patients in all cohorts into high vs low LSC score groups.

The LSC score was tested for associations with survival outcomes as a continuous variable using Cox proportional hazards regression (log-likelihood test) and as a dichotomous stratification (high vs low LSC score) by Kaplan-Meier analysis (log-rank test) using R version 2.11 with survival package 2.35 (R Project for Statistical Computing [http://www.R-project.org]). Patients with missing data were excluded from the analyses. Absolute risk of events occurring by 3 years was determined from Kaplan-Meier analysis. Given that LSCs have been experimentally demonstrated to be resistant to chemotherapy, the LSC score was tested for associations with primary refractoriness to therapy and disease response (event-free and relapse-free survival). For relapse-free survival, only patients who had first achieved clinical remission from disease were included.

The robustness of the association between the LSC score and outcomes was evaluated as follows. The training data set was split in half. Gene weightings defining the LSC score were derived in half of the data and then applied to the other half to test associations with survival. Results from 1000 random splits of the training set were compared. Furthermore, the uniqueness of the prognostic value of the LSC-enriched genes was tested by comparing it with the results obtained by repeating the analysis on 10,000 sets of the same number of randomly selected genes.

Prior clinical investigation in adult AML has defined several important prognostic factors including age, karyotype (chromosomal rearrangements), and molecular mutations, particularly interstitial tandem duplications in the FLT3 gene and mislocalizing mutations in the NPM1 gene. In the current analysis, multivariate Cox regression was used to test whether the LSC score conferred prognostic value independent from these established clinical predictors. Area under the receiver operating characteristic curves were computed using the SurvivalROC package version 1.0 (R Project for Statistical Computing) to further investigate how the LSC score added to known prognostic factors. Because assignments to cytogenetic risk groups were inconsistent between different clinical groups, risk was compared in a uniform fashion across data sets by applying the refined Medical Research Council risk scheme (favorable, intermediate, adverse) based on metaphase karyotypes.

Because acute promyelocytic leukemia is a distinct disease entity, it was excluded from all survival analyses. Association of the LSC score with AML clinical subtypes was also assessed using analysis of variance followed by Games-Howell post hoc testing using SPSS software version 12 (SPSS Inc, Chicago, Illinois) for groups with unequal sizes and variances. All statistical tests were 2-sided. P values of less than .05 were considered significant.

RESULTS

Global gene expression profiles of 15 LSC-enriched subpopulations were compared with 15 paired LPC-enriched subpopulations collected from the same 15 AML samples. The samples were derived from patients representing a diversity of AML subtypes and clinical outcomes (eTable 1 at http://www.jama.com). Significance analysis of microarrays identified 31 genes as more highly expressed in LSC-enriched subpopulations than in LPC-enriched subpopulations. Twenty-one genes were identified as more highly expressed in LPC-enriched subpopulations than in LSC-enriched subpopulations (false discovery rate <10%; Figure 1 and eTable 2). Many of these genes were significantly associated with each other in a network-based analysis (eFigure 2). In addition to the CD34 and CD38 cell surface markers used to purify the samples, the group of genes included factors known to be differentially expressed in early hematopoiesis such as VNN1, RBPMS, SETBP1, GUCY1A3, and MEF2C. All gene symbols reported herein were assigned by the Human Genome Organization Gene Nomenclature Committee. The HOPX homeobox gene was found to be more highly expressed in LSC-enriched subpopulations and has known interactions with the induced pluripotency factors SOX2, POU5F1, NANOG, as well as HDAC2 (eFigure 3).

Gene set enrichment analysis showed that genes more highly expressed in LSCs were enriched for those expressed in normal cells positive for CD34 and negative for CD38, which include normal HSCs compared with normal cells positive for CD34 and CD38 progenitors (Figure 2 and eTable 3). Notably, genes more highly expressed in LSCs also were enriched for genes whose expression in AML has been correlated with high expression of the BAALC gene, which is an adverse prognostic factor. Conversely, proliferation, cell cycle, and differentiation-related genes were systematically repressed in the LSC-enriched subpopulations compared with more differentiated LPC-enriched subpopulations, which is consistent with a tendency for replicative quiescence.

The 31 genes more highly expressed in LSCs were combined to generate an LSC signature (see “Methods” section). The LSC signature was computed in the purified subpopulations from primary AML patient samples and in subpopulations from the normal hierarchy of differentiating myeloid blood cells. By definition, the LSC signature was high in LSCs compared with LPCs, but also relative to downstream more differentiated AML blast cells negative for CD34 (Figure 3). Among normal samples from healthy individuals, the LSC signature was high in HSCs and in multipotent progenitors compared with more mature myeloid cell populations. These observations suggest that the LSC signature is shared with normal HSCs, implying that...
it may reflect self-renewal ability and relative proliferative quiescence.

We next evaluated whether expression of LSC-enriched genes was associated with clinical outcomes using 4 public data sets of bulk AML expression profiles with available clinical annotations. Details of patient characteristics, primary therapies, clinical responses, remission rates, and outcomes have been reported previously16-19,26 and are summarized in Table 1.

The LSC signature was calculated for a training set of 163 patients with AML lacking chromosomal abnormalities (normal karyotype AML).16 The weights were defined for combining expression levels of the 31 LSC genes into a single measure for each sample (eTable 4 at http://www.jama.com). Because the same gene weights were applied to the independent test cohorts, we refer to this as the LSC score. In the training set, the LSC median score was 24.9 (range, 17.4-33.1) and was associated with overall survival as a continuous variable (hazard ratio [HR], 1.15; 95% confidence interval [CI], 1.00-1.30) and was associated with inferior outcome likelihood (P = .002). This association was significant whether patients with normal karyotypes from 1 well-characterized cohort of adult patients with diverse karyotype AML (primarily treated with induction regimens including cytarabine and anthracyclines in Tomasson et al17), the LSC score ranged from 16.6 to 31.0 and was significantly associated with overall survival (Table 2, Figure 4, and Table 3). This association was significant whether the LSC score was evaluated as a continuous predictor (HR, 1.13 [95% CI, 1.04-1.22]; P = .003) or as a high vs low LSC score split (HR, 2.7 [95% CI, 1.4-5.1]; P = .002). Patients in the low LSC score group had a median overall survival of 56.3 months (absolute risk, 39%; 95% CI, 20%-54%) compared with 16.3 months (absolute risk, 81%; 95% CI, 61%-90%) for patients in the high LSC score group (Table 3). The set of genes comprising the LSC score was significant in its prognostic utility when compared with 10 000 randomly selected gene sets of the same

Figure 1. Differences in Gene Expression Between LPC- and LSC-Enriched Subpopulations

<table>
<thead>
<tr>
<th>Gene</th>
<th>Stanford cases12</th>
<th>RIKEN cases13</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOU027883</td>
<td>SETBP1</td>
<td>HLF</td>
</tr>
<tr>
<td>GIMAP7</td>
<td>ABCC2</td>
<td>HOPX</td>
</tr>
<tr>
<td>TMEM500A</td>
<td>PDCD4</td>
<td>C3orf48</td>
</tr>
<tr>
<td>FAM36</td>
<td>RBPS</td>
<td>LTB</td>
</tr>
<tr>
<td>GAOY1A3</td>
<td>SHSBP5</td>
<td>SLC37A3</td>
</tr>
<tr>
<td>SLC37A3</td>
<td>GIMAP6</td>
<td>CD94</td>
</tr>
<tr>
<td>VNN1</td>
<td>MMN1</td>
<td>MMRN1</td>
</tr>
<tr>
<td>SLC38A1</td>
<td>GIMAP2</td>
<td>BIRC3</td>
</tr>
<tr>
<td>LDOC100128550</td>
<td>HSPC159</td>
<td>EVI2A</td>
</tr>
<tr>
<td>PION</td>
<td>FBX02</td>
<td>MEF2C</td>
</tr>
<tr>
<td>ICAM1</td>
<td>HECA</td>
<td>ANLN</td>
</tr>
<tr>
<td>ZWINT</td>
<td>CCL5</td>
<td>C3orf3</td>
</tr>
<tr>
<td>CCL5</td>
<td>C3orf3</td>
<td>UBE2T</td>
</tr>
<tr>
<td>CCL8</td>
<td>KIAA0101</td>
<td>CCNA1</td>
</tr>
<tr>
<td>APOB5</td>
<td>DLGAP5</td>
<td>MSH4</td>
</tr>
<tr>
<td>IL1</td>
<td>OLFM4</td>
<td>Unkown</td>
</tr>
<tr>
<td>ICAM1</td>
<td>UBE2T</td>
<td>MMN1</td>
</tr>
<tr>
<td>BIRC3</td>
<td>STAR</td>
<td>C12TA</td>
</tr>
<tr>
<td>MEB</td>
<td>C12TA</td>
<td>CP47</td>
</tr>
<tr>
<td>P21</td>
<td>HOPX</td>
<td>RASSF3</td>
</tr>
<tr>
<td>SC01</td>
<td>CD93</td>
<td>RASSF2</td>
</tr>
<tr>
<td>CD03</td>
<td>CD08</td>
<td>CD34</td>
</tr>
</tbody>
</table>

Gene expression heat map with each column representing the difference in expression between subpopulations enriched with leukemic stem cells (LSCs) or leukemia progenitor cells (LPCs) isolated from the same patient with acute myeloid leukemia.12-13 Hs indicates LSC/LPC profile purified from primary human patient specimens; Mm, corresponding samples from mouse xenografts. A total of 52 unique genes were identified as differentially expressed between LSC and LPC at a 10% false discovery rate (eTable 2 at http://www.jama.com), with red indicating higher expression in LSC.

©2010 American Medical Association. All rights reserved.
stem cell gene expression and clinical outcomes in acute myeloid leukemia

size (eFigure 6), supporting the conclusion that the association with clinical outcomes was not a false-positive result.

The LSC scores ranged from 16.6 to 35.5 among all patients with non–acute promyelocytic leukemia, including those with cytogenetic abnormalities in the Tomasson et al17 cohort (Table 3), and were associated with overall survival as a continuous variable (HR, 1.10 [95% CI, 1.04-1.17]; P = .001). Patients in the low LSC score group had a median overall survival of 56.3 months (absolute risk, 45%; 95% CI, 30%-57%) compared with 16.5 months (absolute risk, 75%; 95% CI, 64%-83%) in the high LSC score group (HR, 2.0 [95% CI, 1.3-3.2]; P = .003). Investigation of the LSC score in patients from 2 additional cohorts that included patients with chromosomal abnormalities confirmed its association with adverse overall survival (eFigure 5 and Table 2).

Higher LSC scores were consistently associated with inferior event-free survival in patients with normal karyotype AML from 2 cohorts with available data (continuous variable HR, 1.15 [95% CI, 1.06-1.26] P = .001; and continuous variable HR, 1.11 [95% CI, 1.03-1.21] P = .007). The group with the highest LSC scores had a median event-free survival of 10 months (absolute risk, 81%; 95% CI, 60%-91%) compared with 48 months (absolute risk, 48%; 95% CI, 27%-63%) for the group with low LSC scores in the Tomasson et al17 cohort (HR, 2.4 [95% CI, 1.3-4.5]; P = .006) (Figure 4, Table 3, and Table 4). In the second cohort with available event-free survival data,19 the absolute risk for the low LSC score group was 61% (95% CI, 46%-72%) compared with 80% (95% CI, 64%-89%) for the high LSC score group (HR, 1.7 [95% CI, 1.1-2.7] P = .02; eFigure 5).

For the latter data set,10 LSC score also was associated with relapse-free survival in patients with normal karyotype AML who had achieved an initial clinical remission (continuous variable HR, 1.13 [95% CI, 1.01-1.27]; P = .03). The median relapse-free survival was 66 months (absolute risk, 43%; 95% CI 26%-56%) in the cases with low LSC scores compared with only 10 months (absolute risk, 68%; 95% CI, 46%-81%) in the cases with high LSC scores (HR, 1.8 [95% CI, 1.0-3.3] P < .06; eFigure 5 at http://www.jama.com). This finding is consistent with the demonstrated chemoresistance of LSCs.13

Concordantly, the rate of clinical remission was superior among patients with low LSC scores, both in an older cohort (median age of 65 years) with 56% clinical remission for the low LSC score group vs 29% for the high LSC score group (Fisher exact test P < .001)18 and in a younger cohort (median age of 43 years) with 88% clinical remission for the low LSC score group vs 76% for the high LSC score group (Fisher exact test P = .02).20 Furthermore, LSC scores were significantly higher in patients unable to achieve clinical remission compared with those who reached clinical remission (P < .001; Figure 5), a distinction most evident for those patients in whom such remissions were durable.

Figure 2. Enrichment Analysis of Relative Expression Between Leukemic Stem Cells (LSCs) and Leukemic Progenitor Cells (LPCs)

<table>
<thead>
<tr>
<th>Gene Set</th>
<th>Low in LSC vs LPC</th>
<th>High in LSC vs LPC</th>
<th>NES</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Up in CD34+/CD38– vs CD34+/CD38+</td>
<td>2.5</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Up in CD133+ vs CD133–</td>
<td>2.1</td>
<td>.004</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Up in poor outcome (high BAALC) AML</td>
<td>2.1</td>
<td>.005</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proliferation genes</td>
<td>-2.6</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HSC proliferation signature</td>
<td>-2.3</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell cycle genes</td>
<td>-2.2</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Analysis of 17 119 genes (eTable 2 at http://www.jama.com for gene set definitions). Vertical bars in each of the 6 rows represent genes from each of the indicated gene sets. All nominal P values were less than .001. AML indicates acute myeloid leukemia; FDR, false discovery rate; NES, normalized enrichment score.

Figure 3. Expression of the Leukemic Stem Cell (LSC) Signature Across Cell Subpopulations
The LSC score carried prognostic value independent of other known clinical factors. In multivariate Cox regression including age, mutations in FLT3 and NPM1, and cytogenetic abnormalities, the HRs for LSC score in the 3 cohorts were 1.07 (95% CI, 1.01-1.13; \( P = .02 \)), 1.10 (95% CI, 1.03-1.17; \( P = .005 \)), and 1.17 (95% CI, 1.05-1.30; \( P = .005 \)) (eTable 5 at http://www.jama.com). The LSC score was significant independent of these factors in multivariate Cox regression, with high LSC score again being associated with adverse overall survival, event-free survival, and relapse-free survival in all but 1 instance (eTable 5). Comparisons of area under the receiver operator characteristic curves (a measure of the accuracy of a prognostic model) showed that the LSC score added to the prognostic value of age, internal tandem duplications within the FLT3 gene, mislocalizing mutations of the NPM1 gene, and cytogenetic risk in predicting overall survival at 2 years in all cohorts for both normal karyotype AML and AML with chromosomal abnormalities (eTable 6). Models that incorporated the LSC score in addition to other known prognostic factors had consistently higher area under the receiver operator characteristic curves compared with models that did not include the LSC score (eSupplement).

In normal karyotype AML, higher LSC score was associated with inferior overall survival in cases with wild-type NPM1 and mutant NPM1 (despite the fact that the latter are frequently negative for CD34) and in patients with both wild-type FLT3 and wild-type NPM1 (eTable 7). Furthermore, similar results were obtained when all analyses (including derivation

### Table 1. Summary of 4 Independent Acute Myeloid Leukemia (AML) Cohorts With Public Gene Expression Data on Bulk Samples

<table>
<thead>
<tr>
<th>Adult AML Cohort</th>
<th>Metzeler et al(^{16,27})</th>
<th>Wouters et al(^{12,26})</th>
<th>Tomasson et al(^{17,20})</th>
<th>Wilson et al(^{18})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Karotype</td>
<td>Normal</td>
<td>Mixed</td>
<td>Mixed</td>
<td>Mixed</td>
</tr>
<tr>
<td>Age, median (IQR), y</td>
<td>60 (17-85)</td>
<td>46 (15-77)</td>
<td>47 (16-81)</td>
<td>65 (20-84)</td>
</tr>
<tr>
<td>No. of patients</td>
<td>163</td>
<td>526</td>
<td>188</td>
<td>170</td>
</tr>
<tr>
<td>Follow-up, median (IQR), mo</td>
<td>9 (3-27)</td>
<td>17 (7-73)</td>
<td>25 (10-55)</td>
<td>10 (2-22)</td>
</tr>
<tr>
<td>Cooperative group</td>
<td>German AML Cooperative Group</td>
<td>Dutch-Belgian Hematology-Oncology Cooperative</td>
<td>Washington University(^a) and CALGB</td>
<td>Southwest Oncology Group</td>
</tr>
<tr>
<td>Patient outcome data</td>
<td>Overall survival</td>
<td>Response to therapy; overall, event-free, and relapse-free survival</td>
<td>Overall and event-free survival</td>
<td>Response to therapy; overall survival</td>
</tr>
<tr>
<td>Affymetrix microarray platform</td>
<td>HG-U133 A&amp;G</td>
<td>HG-U133 Plus 2.0</td>
<td>HG-U133 Plus 2.0</td>
<td>HG-U95Av2</td>
</tr>
<tr>
<td>Data set accession No.</td>
<td>GSE12417</td>
<td>GSE14488</td>
<td>GSE10358</td>
<td>wilm-00119(^c)</td>
</tr>
</tbody>
</table>

Abbreviations: CALGB, Cancer and Leukemia Group B; GSE, Gene Expression Omnibus of the National Center for Biotechnology Information; IQR, interquartile range.

\(^{a}\)Reflect earliest patient enrollment to most recent follow-up as captured by clinical trial registration, enrollment, and publication dates.

\(^{b}\)National Cancer Institute caArray database (https://array.nci.nih.gov/caarray/project/wilm-00119).

<table>
<thead>
<tr>
<th>Test Cohorts</th>
<th>Wouters et al(^{12,26})</th>
<th>Tomasson et al(^{17,20})</th>
<th>Wilson et al(^{18})</th>
<th>Metzeler et al(^{16,27})</th>
</tr>
</thead>
<tbody>
<tr>
<td>NKAML</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSC score</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continuous(^b)</td>
<td>1.17 (1.07-1.28)</td>
<td>1.13 (1.04-1.22)</td>
<td>1.18 (1.04-1.34)</td>
<td>1.97 (1.88-2.07)</td>
</tr>
<tr>
<td>Dichotomous(^c)</td>
<td>1.86 (1.15-3.02)</td>
<td>2.70 (1.43-5.10)</td>
<td>2.55 (1.44-2.81)</td>
<td>1.74 (1.37-2.19)</td>
</tr>
<tr>
<td>FLT3 ITD(^d)</td>
<td>1.84 (1.14-2.98)</td>
<td>2.68 (1.42-5.07)</td>
<td>1.28 (0.73-2.35)</td>
<td>2.22 (1.49-3.31)</td>
</tr>
<tr>
<td>NPM1c(^d)</td>
<td>0.76 (0.47-1.23)</td>
<td>1.55 (0.86-2.79)</td>
<td>0.67 (0.38-1.17)</td>
<td>0.79 (0.54-1.17)</td>
</tr>
<tr>
<td>Age, y</td>
<td>1.01 (0.98-1.03)</td>
<td>1.02 (1.00-1.04)</td>
<td>1.03 (1.00-1.06)</td>
<td>1.03 (1.01-1.04)</td>
</tr>
</tbody>
</table>

| Non-APL      |                             |                             |                             |                             |
| LSC score    |                             |                             |                             |                             |
| Continuous\(^b\) | 1.07 (1.02-1.13)            | 1.10 (1.04-1.17)            | 1.15 (1.07-1.25)            | 1.09 (1.01-1.17)            |
| Dichotomous\(^c\) | 1.36 (0.98-1.88)            | 2.01 (1.27-3.18)            | 1.99 (1.43-2.79)            | 1.48 (1.21-1.83)            |
| FLT3 ITD\(^d\) | 1.82 (1.28-2.59)            | 1.82 (1.09-3.02)            | 1.12 (0.78-1.62)            | 1.53 (1.19-2.00)            |
| NPM1c\(^d\) | 0.86 (0.60-1.22)            | 1.49 (0.95-2.32)            | 0.79 (0.55-1.14)            | 0.20 (0.13-0.32)            |
| Age, y       | 1.01 (1.00-1.03)            | 1.03 (1.01-1.04)            | <1.001                      | 1.03 (1.01-1.04)            |
| Cytogenetic risk group\(^g\) | 2.04 (1.54-2.69)         | 1.97 (1.34-2.90)            | 2.16 (1.52-3.06)            | 2.04 (1.54-2.69)            |

Abbreviations: APL, acute promyelocytic leukemia; CI, confidence interval; HR, hazard ratio; ITD, internal tandem duplication; NKAML, normal karyotype AML.

\(^{a}\)Patients with APL were excluded. \(^{b}\)P values were computed using the log-likelihood test for continuous values and the log-rank test for discrete variables. \(^{c}\)Unit of measure is log, of gene expression intensity. \(^{d}\)High vs low score. \(^{e}\)Mutated vs wild type. \(^{f}\)Favorable, intermediate, or adverse per the modified Medical Research Council scheme.\(^{23}\)

©2010 American Medical Association. All rights reserved. (Reprinted) JAMA, December 22/29, 2010—Vol 304, No. 24 2711
Stratification of outcomes using this approach is depicted for overall survival in the training set\(^\text{16}\) and in one of the validation sets and event-free survival in one of the validation sets.\(^\text{17}\) Vertical ticks on curves indicate censored events. Similar results were obtained in the other independent data sets (Table 2 and eFigure 4 at http://www.jama.com).

![Figure 4. Association Between Leukemic Stem Cell (LSC) Score and Survival Outcomes in Patients With Normal Karyotype Acute Myeloid Leukemia (AML)](image)

**Table 3. Range of Leukemic Stem Cell (LSC) Scores Across Acute Myeloid Leukemia (AML) Cohorts With Survival Analyses**

<table>
<thead>
<tr>
<th>AML Cohort</th>
<th>No. of Patients</th>
<th>LSC Score, Median (IQR)</th>
<th>Median Survival by LSC Score Group, mo</th>
<th>Comparative Absolute Risk of Event by LSC Score Group at 3 y, % (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Low</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Low (24.9-27.0)</td>
<td>22.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>High (25.2-27.6)</td>
<td>16.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Low (25.0-27.0)</td>
<td>22.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>High (25.0-27.0)</td>
<td>16.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Low (14.0-15.6)</td>
<td>22.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>High (14.0-15.6)</td>
<td>16.5</td>
</tr>
</tbody>
</table>

Abbreviations: APL, acute promyelocytic leukemia; CI, confidence interval; IQR, interquartile range.

\(^{a}\)The corresponding hazard ratios and \(P\) values appear in Table 2 and Table 4.

\(^{b}\)The primary end point was overall survival unless otherwise indicated.

\(^{c}\)Data set does not have probes for some of the LSC genes, hence the range is different from the other 3 cohorts.

Typically, LSC score gene weightings were performed after excluding cases with cells negative for CD34 (defined either as having mutant NPM1 or as the 40% of samples with the lowest CD34 expression). Exclusion of cases with cells negative for CD34 from the model and validation resulted in an LSC score with similar gene weightings and prognostic value (eTable 4). Therefore, the LSC score was not simply a proxy for CD34 status. Taken together, these data indicate that higher LSC score is associated with inferior survival outcomes independent of age, internal tandem duplications within FLT3 or NPM1 mutations, CD34 expression, and cytogenetic risk group and adds to their prognostic utility.

Although similar across most age groups and morphological subtypes, LSC scores were higher in cases with minimally differentiated myeloblasts (French-American British M0), which typically have poor prognosis,\(^\text{28}\) consistent with previous reports of high LSC prevalence in this subtype (eFigure 7).\(^\text{29}\) In general, the LSC score was similar in favorable, intermediate, and adverse cytogenetic risk groups, and was not a direct proxy for this factor (eFigure 7C). This is consistent with our findings from survival analyses that showed that the LSC score confers an independent prognostic value. When considering specific cytogenetic subgroups, the LSC score had higher than average values in patients with unfavorable −5 or 7(q) abnormalities, and lower than average values among AML-harboring anomalies involving 11q23/MLL. Recent studies have reported that self-renewing cells from AML mouse models carrying MLL anomalies reside in more mature cells.\(^\text{30}\)

We also investigated the relationship of the LSC score with molecular mutations in normal karyotype AML, which is the largest single cytogenetic subgroup of AML. The LSC scores were significantly lower in those harboring mislocalizing mutations of the NPM1 gene (eFigure 7D and eFigure 8), in agreement with recent observations that leukemia initiating cells in NPM1-mutant AML are frequently negative for FLT3 and mislocalizing mutations of the...
NPM1 gene, which is a combination
confering a distinctly favorable prog-
nosis in patients with normal karyo-
type AML (eFigure 7D).8 The LSC
scores also were lower in patients with
normal karyotype AML and double
CEBPA mutations, which is again as-
associated with favorable outcomes19 rela-
tive to cases with single mutations but
not relative to wild-type CEBPA. Simi-
lar findings were observed in all 4 in-
dependent data sets totaling 1047 pa-
tients (eFigure 8 and eFigure 9 at http:
//www.jama.com). Of note, no
significant differences in LSC scores
were observed when patients with AML
were stratified according to less com-
mon recurrent somatic mutations, in-
cluding those in the tyrosine kinase do-
main of FLT3 or activating mutations
in NRAS, KRAS, or IDH1.

**COMMENT**

Clinical evidence supporting the signifi-
cance of the cancer stem cell model for
human AML has been lacking despite
ample experimental evidence from trans-
plantation assays in immunocompro-
mised mice. In this study, we show that
a gene expression score associated with
the LSC-enriched subpopulation is an
independent prognostic factor in AML,
with high LSC score associated with ad-
verse outcomes in multiple indepen-
dent cohorts. Specifically, high LSC score
is associated with poor overall sur-
vival, event-free survival, and relapse-
free survival in patients with normal karyo-
type AML and inferior overall sur-
vival in patients with chromosomal ab-
normalities. Additionally, the LSC score
was associated with primary response to
induction chemotherapy because high
LSC scores strongly correlated with
lower remission rates. Multivariate
analysis demonstrated that high LSC
score was associated with poor out-
comes independently of age, presence of
FLT3 or NPM1 mutations, and cytoge-
netic risk group. These findings sup-
port the clinical relevance of the cancer
stem cell model for AML.

Stem cells of AML were originally
identified by prospectively separating pri-
mary leukemic specimens into subpopu-
lations based on expression of CD34 and
CD38, which are surface markers that are
differentially expressed in normal hem-
atopoiesis (eFigure 1).10 When the
function of these tumor subpopula-
tions was assessed by transplantation into
immunodeficient mice, leukemia-
initiating activity was demonstrated exclu-
sively in the fraction of cells positive for
CD34 and negative for CD38.11 The
majority of recent studies indicate that
LSC activity for AML is enriched in the
subpopulation of cells positive for CD34
and negative for CD38, although recent
reports have challenged whether this is
exclusive.12,31,32 The clinical significance
of the LSC model is suggested by 2 prior
studies, the first of which identified an
inverse correlation between the fre-
quency of cells positive for CD34 and
negative for CD38 at diagnosis and the
duration of relapse-free survival.33 The
second study reported that the relative
ability of AML cells to successfully en-
graft in immunodeficient mice (a prop-
erty associated with LSC) correlated with
adverse clinical features.34 While sug-
gestive, neither of these studies investi-
gated large cohorts of patients with long-
term follow-up and diverse clinical
features.

Notably, the LSC signature was highly
expressed in purified HSC and

### Table 4. Leukemic Stem Cell (LSC) Score as a Univariate Predictor of Event- and Relapse-Free Survival in 2 Independent Acute Myeloid Leukemia (AML) Cohorts

<table>
<thead>
<tr>
<th></th>
<th>Wouters et al16,18</th>
<th>Tomasson et al17,20</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Event-Free Survival</td>
<td>Relapse-Free Survival</td>
</tr>
<tr>
<td></td>
<td>HR (95% CI)</td>
<td>P Value</td>
</tr>
<tr>
<td>NKAML</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSC score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continuousb</td>
<td>1.15 (1.06-1.26)</td>
<td>.001</td>
</tr>
<tr>
<td>Dichotomousc</td>
<td>1.78 (0.99-3.25)</td>
<td>.06</td>
</tr>
<tr>
<td>FLT3 ITDd</td>
<td>1.09 (0.59-2.02)</td>
<td>.79</td>
</tr>
<tr>
<td>NPM1c</td>
<td>0.92 (0.58-1.47)</td>
<td>.74</td>
</tr>
<tr>
<td>Age, y</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.99 (0.97-1.02)</td>
<td>.63</td>
</tr>
</tbody>
</table>

Abbreviations: APL, acute promyelocytic leukemia; CI, confidence interval; HR, hazard ratio; ITD, internal tandem du-
plication; NKAML, normal karyotype AML.

a Patients with APL were excluded. P values were computed using the log-likelihood test for continuous values and the
log-rank test for discrete variables.

b Unit of measure is log2 of gene expression intensity.

c High vs low score.

d Mutated vs wild type.

The LSC score was significantly associated with initial therapeu-
ic response as determined by the ability to
achieve clinical remission in 2 data sets for which this
information was available.13,14 Boxes indicate the inter-
quartile range with median shown as the thick horizon-
tal bar. T test P < .001 for comparison of clinical remis-
sion with no remission for study by Wouters et al18 and
Wilson et al19 Y-axis shown in blue indicates range of
LSC score from 10 to 20. Error bars indicate full range.

**Figure 5. Association Between Leukemic Stem Cell (LSC) Score and Initial Therapeutic Response**

©2010 American Medical Association. All rights reserved.

(Reprinted) JAMA, December 22/29, 2010—Vol 304, No. 24 2713
was much less expressed in more differentiated myeloid progenitor cells, suggesting that it may be reflective of self-renewal ability. Despite the observed similarities between the LSC signature and HSC gene expression programs, therapeutic targeting of LSCs is still possible without toxicity toward normal HSC. Indeed, markers distinguishing LSC from HSC exist and are amenable to targeted therapies, including antibodies to CD47, CLL-1, and CD123.42,43 Future work is needed to prospectively validate the prognostic ability of the LSC score by evaluating its component genes using reverse transcription–polymerase chain reaction in an independent patient cohort. It will also be pertinent to examine the relationship of the LSC score to other gene expression signatures that have been proposed for predicting survival in patients with AML.16,26,38

In addition to the cell markers of CD34 and CD38 that were used for their purification, LSCs were distinguished from LPCs by the expression of several other genes known to be differentially expressed during early myelopoiesis. These included 3 members (GIMAP2, GIMAP6, and GIMAP7) of a small family of immunoassociating nucleotide-binding proteins implicated in survival of HSCs and leukemia.39 However, no prior associations with AML have been described. Two genes (HOPX and GUCY1A3) in this signature, which have previously been incorporated into AML prognostic models,16,38 are notable for their distinctive pattern of expression and histone modification in self-renewing cells.40 HOPX is an unusual homeomain protein known to directly recruit histone deacetylase activity without directly binding DNA.41 This gene also is directly repressed in vivo in malignant cells in response to administration of the histone deacetylase inhibitor panobinostat.42 The latter is currently being studied in clinical trials for patients with AML. GUCY1A3, which encodes a component of the soluble guanylate cyclase enzyme catalyzing the conversion of guanosine triphosphate to cyclic guanosine monophosphate, is repressed during replicative senescence.43 Cyclic guanosine monophosphate has been reported to stimulate HSC proliferation.44

The cancer stem cell model has been studied in solid tumors in addition to leukemia.7 Investigation of gene expression in human breast cancer stem cells identified a signature prognostic of metastasis-free survival and overall survival in multiple carcinomas, suggesting the clinical significance of the cancer stem cell model in these solid tumors.45 Among other human malignancies, we and others have described the prognostic significance of distinctive signatures of self-renewing populations, including embryonic stem cells,46,47 HSCs, and progenitor cells,48,49 our study is the first to directly define a signature of enriched AML-initiating cells and to relate this signature to expression profiles of diagnostic specimens, allowing a link to corresponding clinical and pathological features of patients. Ultimately, this model has major implications for cancer therapy, most notably that in order to achieve cure, the cancer stem cells must be eliminated.7 To accomplish this in AML, novel therapies targeting LSC must be developed. Several therapies are being investigated including small molecules21,50-52 and monoclonal antibodies,53,56,57 which hold promise for improving therapeutic efficacy beyond current conventional treatments.

Our LSC prognostic model requires validation in a prospective study of patients with AML treated with a standardized protocol incorporating strict eligibility criteria, a uniform treatment plan, uniform sample collection and handling, and well-defined primary end points. Moreover, microarrays are not broadly used in clinical decision making.54 While surrogate methods such as real-time polymerase chain reaction have demonstrated clinical utility,55 their application requires performance assessment in independent laboratories. Flow cytometric analysis of the predictive power of the proteins comprising the LSC score and comparison with RNA-based models may help to determine the best platform for future clinical application. However, the monoclonal antibodies useful for flow cytometry are not presently available for the full set of encoded proteins.

**CONCLUSION**

High expression of an LSC gene signature is independently associated with adverse outcomes in patients with AML. If prospectively validated, the described LSC score may be incorporated into routine clinical practice for predicting prognosis in patients with AML and used in clinical trials incorporating risk-based stratification or randomization strategies.

**Author Contributions:** Drs Gentles and Alizadeh had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: Gentles, Plevritis, Majeti, Alizadeh. Acquisition of data: Gentles, Majeti, Alizadeh. Analysis and interpretation of data: Gentles, Plevritis, Majeti, Alizadeh. Drafting of the manuscript: Gentles, Plevritis, Majeti, Alizadeh. Critical revision of the manuscript for important intellectual content: Gentles, Plevritis, Majeti, Alizadeh. Statistical analysis: Gentles, Plevritis, Alizadeh. Obtained funding: Plevritis, Majeti. Administrative, technical, or material support: Plevritis, Majeti, Alizadeh. Study supervision: Majeti, Alizadeh.

**Financial Disclosures:** A patent application for the development of the leukemic stem cell score as a diagnostic assay has been submitted by Drs Gentles, Majeti, and Alizadeh. Dr Plevritis did not report any financial disclosures.

**Funding/Support:** This research was supported by the Integrative Cancer Biology Program through National Institutes of Health awards 1U54CA149145 and U56-CA112973 (Dr Plevritis). Dr Majeti holds a Career Award for Medical Scientists from the Burroughs Wellcome Fund. Dr Alizadeh holds a Career Development Award from the Leukemia and Lymphoma Society.

**Role of the Sponsors:** The funding organizations had no role in the design and conduct of the study; in the collection, analysis, and interpretation of the data; or in the preparation, review, or approval of the manuscript.

**Online-Only Material:** The eSupplement, eTables 1-7, and eFigures 1-9 are available at http://www.jama.com.

**Additional Contributions:** We thank Irving Weissman, MD, and Ronald Levy, MD, for support and discussion and Rob Tibshirani, PhD, for advice on statistical analyses (all 3 from Stanford University, Palo Alto, California). We also thank the following Stanford University clinicians for critically reading the manuscript: Bruno Medeiros, MD, Jason Gotlib, MD, Linda Boxer, MD, PhD, and Beverly Mitchell, MD. We are indebted to the patients and their physicians, including the Stanford Hematology Tissue Bank, the German AML Cooperative Group, the Haematolog-
Oncology Foundation for Adults in the Netherlands, the European Group, the German AML Study Group/Ulm, Washington University School of Medicine, and the Cancer and Leukemia Group B, for sharing raw data and clinical information from their studies. The Austrian Acute Myeloid Leukemia Working Parties. The importance of diagnostic cytogenetic data in AML. Blood. 1998;92(7):2322-2333.


