Relative Value of ZAP-70, CD38, And Immunoglobulin Mutation Status In Predicting Aggressive Disease In Chronic Lymphocytic Leukemia

Laura Z. Rassenti¹, Sonia Jain², Michael J. Keating⁴, William G. Wierda⁴, Michael R. Grever⁷, John C. Byrd⁷, Neil E. Kay⁸, Jennifer R. Brown³, John G. Gribben⁵, Donna S. Neuberg³, Feng He², Andrew W. Greaves², Kanti R. Rai⁶, Thomas J. Kipps¹†

Division of Hematology/Oncology, Department of Medicine, University of California, San Diego, La Jolla, CA; Moores UCSD Cancer Center, La Jolla, CA; Division of Biostatistics and Bioinformatics, University of California, San Diego, CA; Dana Farber Cancer Institute, Boston, MA; Department of Leukemia, the University of Texas M.D. Anderson Cancer Center, Houston, TX; Barts and The London School of Medicine, London, UK; Long Island Jewish Medical, New Hyde Park, NY; James Cancer Institute, Ohio State University, Columbus, OH; Department of Hematology, Mayo Clinic, Rochester, MN, The CLL Research Consortium.

†To whom correspondence should be addressed: (e-mail tkipps@ucsd.edu, ph (858) 534-5417)

*Funded in part by NIH grant PO1-CA081534 for the CLL Research Consortium (CRC)
ABSTRACT

Leukemia-cell expression of ZAP-70, CD38, or unmutated immunoglobulin heavy-chain-variable region genes (U-IGHV) each is associated with aggressive disease in patients with chronic lymphocytic leukemia (CLL). To assess the relative strength of each marker we defined thresholds for designating a case as positive for CD38 or ZAP-70 in a test cohort of 307 patients and used these data-defined criteria to stratify patients in an independent cohort of 705 patients. Multivariable analysis revealed that ZAP-70 was the strongest risk factor. Knowledge of the IGHV mutation status or CD38 did not improve our ability to predict the time to first treatment except for ZAP-70-negative cases, which could be segregated into two groups of intermediate-risk or low-risk disease based upon whether they expressed unmutated or mutated IGHV. ZAP-70 maintained its high relative prognostic value for the subset of patients with early-stage, asymptomatic disease, including patients evaluated within one year of diagnosis. Although it is premature to recommend therapy based on these risk factors, patients with ZAP-70-positive CLL cells should be monitored closely for disease progression as they have a median time from diagnosis to requiring initial therapy by standard criteria of approximately three years.
INTRODUCTION

The clinical behavior of patients with chronic lymphocytic leukemia (CLL) is heterogeneous. Whereas some patients have indolent disease and lack disease-related complications for many years, others develop progressive and/or symptomatic disease requiring therapy within a relatively short time after diagnosis. Early treatment of the former could place patients at risk for therapy-related complications that might compromise their quality of life and/or survival\(^1,2\). Defining markers that reliably can stratify patients into groups with good-risk or poor-risk disease could facilitate clinical trials evaluating the potential benefit of early treatment.

Several markers each can segregate patients into subgroups with different rates of disease progression. The expression of unmutated immunoglobulin heavy chain variable region genes (IGHV) predicts more aggressive clinical behavior\(^3\)-\(^8\), as does leukemia-cell expression of CD38\(^3\),\(^7\)-\(^17\). Gene expression analyses found that CLL cells with unmutated IGHV (U-IGHV) differed from CLL cells with mutated IGHV (M-IGHV) in the expression levels of a relatively small subset of genes, one of which encodes the zeta-chain associated protein of 70kD (ZAP-70), an intracellular tyrosine kinase involved in T cell-receptor signaling\(^18,19\). Measurements of this intracellular protein can be used as a surrogate marker for expression of U-IGHV\(^20\)-\(^25,38,39\).

We used recursive partitioning on flow cytometry and clinical data to define the optimal threshold for designating a sample as CD38-positive among samples collected from a test cohort of 307 patients with CLL who previously were characterized for expression of ZAP-70 and IGHV\(^22\). Using data-defined criteria for designating a sample positive for CD38 or ZAP-70, we assessed the relative value of CD38, ZAP-70, or use
of U-IGHV in predicting the time to initial treatment of patients in an independent validation cohort of 705 CLL patients followed by the CLL Research Consortium (CRC).
MATERIALS AND METHODS

Patients Samples And Cell-Processing

Blood was obtained from CLL patients enrolled in the CRC upon written informed consent. Blood mononuclear cells were prepared using Ficoll-Hypaque 1077 (Sigma, St. Louis, MO) and suspended in fetal-calf serum containing 10% dimethylsulfoxide for storage in liquid nitrogen\textsuperscript{22}. The IRB of UCSD has reviewed and approved this study in accordance with the requirements of the Code of Federal Regulations on the Protection of Human Subjects and the Declaration of Helsinki.

Patient Characteristics

The test cohort was comprised of 307 CLL patients (201 males (65%) and 106 females (35%)) who had a median age at diagnosis of 52 years (range 30-77 years). The validation cohort was comprised of 705 CLL patients (466 males (66%) and 239 females (34%)) who had a median age at diagnosis of 56 years (range 27-83 years). The Rai stage was assessed when the blood sample was collected. The distribution of patients with modified Rai stage (e.g. low-risk (stage 0), intermediate-risk (stages I and II), or high-risk (stages III and IV) disease was 20%, 50%, or 30%, respectively, in the test cohort, and 22%, 47%, or 31%, respectively, in the validation cohort. The median follow-up period from sample collection to last follow-up was 5.4 years.

Patients received treatment when they developed symptomatic and/or progressive disease, as per NCI-working-group criteria\textsuperscript{26}. Patients confirmed to have remained untreated were censored for this endpoint. Eighty-four patients in the test cohort (27%) and 184 patients in the validation cohort (26%) received treatment for CLL.
before sample collection. For these patients, the median time from diagnosis to initial
treatment was 1.8 years and 1.5 years for the test and the validation cohorts,
respectively. Sixty-nine of the 223 untreated patients in the test cohort (31%) and 172
of the 521 untreated patients in the validation cohort (33%) received treatment after the
sample was collected. For these patients, the median time from diagnosis to first
treatment was 4.0 years and 3.6 years, respectively. Collectively, a total of 356 of the
705 patients in the validation cohort (50%) received therapy for CLL, of which 184 of
705 were treated prior sample collection and 172 of 705 were treated after sample
collection.

**Clinical Database**

All the clinical data was captured via the Tissue-Core-Management-System (TCMS).
We used closed data fields to assess the date of diagnosis, date of initial treatment,
and clinical status of each patient. A multiple closed unique patient identifier system
was used that complied with the Health Insurance Portability and Accountability Act.

**Assays For ZAP-70 And CD38**

We performed flow cytometry to detect expression of ZAP-70 on cryopreserved
samples of each patient in the validation cohort (N=705). CLL cells also were
analyzed for CD19, CD20, CD23, and CD38, using mAbs conjugated to
allophycocyanin (APC), Peridinin-chlorophyll-protein (PerCp), fluorescein
isothiocyanate (FITC), or phycoerythrin (PE), respectively (Becton Dickinson,
Pharmingen, La Jolla, CA). Fluorochrome-conjugated, isotype control mAbs of
irrelevant specificity were used in all experiments to monitor for non-specific staining. A case was considered ZAP-70-positive if the percentage of CLL cells expressing ZAP-70 was greater than 20%.

Assessing IGHV Mutation Status And Genomic Alterations

We analyzed the IGHV expressed by the CLL cells of each member of the validation cohort. Sequences with less than 98% homology with the corresponding germ line IGHV were considered mutated. Fluorescent in situ hybridization (FISH) was performed on interphase nuclei of blood lymphocytes. The FISH panel (Vysis Inc, Downers Grove, IL) detected aberrations at 11q22, 12 centromere, 13q14.3, and 17p13. FISH data were available for 19% (58/307) or 46% (323/705) of the patients in the test or validation cohort, respectively.

Statistical Analysis

The time from diagnosis to initial treatment was assessed by Kaplan-Meier curves and the log-rank test. Recursive partitioning identified the optimal threshold for CD38 in the test cohort. Associations among IGHV mutation status or expression of ZAP-70 or CD38 and the time from diagnosis to initial therapy was investigated using the Cox Proportional Hazards Regression Model (Breslow method for ties), the Fisher’s exact test for dichotomous data, the Kruskal-Wallis test for ordered categorical data, and the Wilcoxon rank sum test for continuous data. All p-values were two-sided. Statistical analyses were performed in R (version 2.3.1).
RESULTS

Defining The Threshold For CD38 In The Test Cohort

We determined the optimal cut-off for designating a CLL sample as CD38-positive that best could segregate patients in the test cohort of 307 patients who had short versus long median times to treatment, using recursive partitioning and the log-rank test on flow cytometry and clinical data similar to how we defined the ZAP-70 threshold with this same cohort\textsuperscript{22}. The optimal threshold was when 34\% or more of the CLL cells had fluorescence above the background-fluorescence threshold. Using this threshold, patients in the test cohort who had CLL cells that were classified as being CD38-negative had a median time to first treatment of 7.8 years, whereas patients with CD38-positive CLL cells had a median time to first treatment of 3.4 years (p<0.001).

Prognostic Value Of CD38, ZAP-70, And IGHV Mutation Status In The Validation Cohort

We applied these data-defined criteria to classify CLL-cell samples in a validation cohort of 705 patients. When the threshold of CD38 (greater or equal to 34\%) was applied to stratify patients in this cohort, the 374 patients with CLL cells that were CD38-negative had a median time to first treatment of 5.7 years, whereas the 333 patients with CD38-positive CLL cells had a median time to first treatment of 4.0 years (p<0.001) (Table 1). On the other hand, 331 patients in the validation who had ZAP-70-negative CLL cells had a median time from diagnosis to initial therapy of 8.4 years, whereas the 374 patients with CLL cells found ZAP-70 positive had a median time from diagnosis to therapy of 2.6 years (p<0.001). Similarly, 251 patients in the validation cohort were found to use mutated IGHV (M-IGHV). These patients had a significantly
longer time to first treatment than did 454 patients with CLL cells in the validation cohort found to use unmutated IGHV (U-IGHV) (namely 8.6 years versus 3.8 years, p<0.001) (Table 1). In univariate analysis, the Cox proportional hazard ratio (HR) associated with expression of ZAP-70 for the entire validation cohort was 4.08 (p<0.001), the HR associated with U-IGHV was 3.14 (p<0.001), and the HR associated with expression of CD38 was 1.53 (p<0.001) (Table 2). These HRs were similar for the subset of 521 patients in the validation cohort who enrolled into the CRC prior to initial therapy (Table 2). The HR associated with the expression of ZAP-70 was 5.36 (p<0.001), the HR associated with use of U-IGHV was 3.76 (p<0.001), and the HR associated with the expression of CD38 was 1.51 (p<0.01).

Relative Prognostic Value Of CD38, ZAP-70, And IGHV Mutation Status

We examined the time from diagnosis to initial therapy of patients segregated by their leukemia-cell expression of ZAP-70 into subgroups that then were segregated by CLL-cell expression of CD38 (Figure 1, panels A and B for ZAP-70+ cases or panels E and F for ZAP-70-negative cases) or IGHV mutation status (Figure 1, panels C and D for ZAP-70+ cases, or panels G and H for ZAP-70-negative cases), and then subsequently stratified by IGHV mutation status or expression of CD38, respectively. For the subgroup of patients in the validation cohort who had ZAP-70-positive CLL cells, knowledge of CD38 expression or IGHV mutation status did not change the prediction that these patients were at a higher risk for early treatment (Figure 1, Panels A, B, C, D; p values are not significant). On the other hand, the 331 patients in the validation cohort who had CLL cells that did not express ZAP-70 could be segregated into two
groups with an intermediate or highly indolent clinical course by their IGHV mutation status, irregardless of whether or not the CLL cells expresed CD38 (Figure 1, Panels E and F, p<0.01). However, among these patients with CLL cells that lacked expression of ZAP-70, there were no apparent differences in subgroups that differed in their CLL-cell expression of CD38 when these patients were segregated by IGHV mutation status (Figure 2, panels G and H, p>0.5). Collectively, we observed that knowledge of CLL-cell expression of CD38 did not provide additional prognostic information when the CLL-cell expression-level of ZAP-70 or IGHV mutation status was known (Figure 1, Panels C, D, G, and H).

Similarly, we examined the time from diagnosis to initial therapy of patients segregated by IGHV mutation status into subgroups that were segregated by CLL-cell expression of ZAP-70 (Figure 2, panels A and B for cases that used M-IGHV, or panels E and F for cases that use U-IGHV) or CD38 (Figure 2, panels C and D for for cases that used M-IGHV, or panels G and H for cases that use U-IGHV), and then subsequently stratified by expression of CD38 or ZAP-70, respectively. Again knowledge of CLL-cell expression of CD38 did not provide additional prognostic information when the CLL-cell expression-level of ZAP-70 or IGHV mutation status was known (Figure 2, Panels A, B, E, and F). On the other hand, if the IGHV mutation status was known, testing for ZAP-70 status still contributed significantly to our ability to segregate patients into aggressive versus more indolent subgroups (Figure 2, Panels C, D, G, H, p<0.0001 for each comparison).

In a multivariable analysis, the Cox proportional HR associated with expression of ZAP-70 (4.08) was not substantially altered when either the expression of U-IGHV or
CD38 was added to the model (3.09 and 3.96, respectively) (Table 2). Likewise, the HR associated with expression of ZAP-70 was significant when both expression of U-IGHV and CD38 were added to the model (HR=3.01, p<0.001) (Table 2). However, the HR associated with expression of U-IGHV (HR=3.14) was reduced to 1.87 when ZAP-70 was added to the model, but not altered when CD38 was added to the model (HR=3.05) (Table 2). The low HR associated with CD38 (HR=1.53) became trivial when ZAP-70 was added to the model (HR=1.14, P=0.25) (Table 2) and was lowered slightly when IGHV mutation status was added to the model (HR=1.39, p=0.002) (Table 2).

A similar hierarchy of HRs was noted for the subset of 152 patients in the validation cohort who were over 65 years of age at diagnosis (Table 2). These patients had a median age at diagnosis of 70, which is comparable to the median age of all patients diagnosed with CLL28. For these patients the HR associated with expression of ZAP-70 (2.79) was not substantially altered when either the expression of U-IGHV or CD38 was added to the model (2.41 and 2.84, respectively) (Table 2). Likewise, the HR associated with expression of ZAP-70 was significant when both expression of CD38 and U-IGHV were added to the model (HR=2.47, p<0.001) (Table 2). However, the HR associated with expression of U-IGHV (HR=2.14) was reduced to 1.43 when ZAP-70 was added to the model, but not altered when CD38 was added to the model (HR=2.15) (Table 2). The low HR associated with CD38 (HR=1.08) remained trivial when ZAP-70 or IGHV mutation status was added to the model (Table 2). This analysis revealed that the relative strengths of each of these markers in predicting the need for early therapy noted in the entire validation cohort also applied for the subgroup
of patients in this cohort who had a median age at diagnosis similar to that of all patients diagnosed with this disease in the United States\textsuperscript{28}.

In multivariable analysis, the relative strength of CLL-cell expression of ZAP-70 in predicting early treatment also was apparent for the 521 patients in the validation cohort who had blood samples collected prior to initial therapy. Moreover, in this subset of patients, the HR associated with CLL-cell expression of ZAP-70 appeared highest for patients with low-risk or high-risk Rai-stage disease, but comparable to the HR associated with use of U-IGHV for patients with intermediate-risk Rai-stage disease (Table 2). However, the HR associated with CLL-cell expression of CD38 was the lowest in all subgroups, regardless of stage (Table 2).

We also performed an analysis on the 199 patients in the validation cohort who presented to a CRC physician with early-to-intermediate Rai-stage disease (Rai stages 0-2) within one year of diagnosis before receiving anti-leukemia therapy. The median time from diagnosis until when these patients contributed samples for this study was 0.24 years and the median follow-up period from sample collection to last follow-up was 3.0 years. At the time of these analyses, 15 (15\%) of the 97 patients in this subgroup who had ZAP-70-negative CLL cells, had received treatment, whereas 48 (47\%) of the 102 patients with ZAP-70-positive CLL cells had received treatment (p<0.001). On the other hand 34 (30\%) of the 113 patients with CLL cells found negative for CD38 had received treatment, whereas 29 (34\%) of the 86 patients with CLL cells considered CD38-positive had received treatment (p=0.60). Of the 70 patients with CLL cells that used M-IGHV, 14 (20\%) had received treatment, whereas 49 (38\%) of the 129 patients with CLL cells that expressed U-IGHV had received treatment (p=0.01). Collectively,
the HR associated with CLL-cell expression of ZAP-70 was 4.85 (P<0.001), the HR associated with use of U-IGHV was 2.26 (P<0.01), and the HR associated with expression of CD38 was only 1.26 (P=0.35) (Table 2).

**Relationship Between CLL-cell Expression of ZAP-70 And IGHV Mutation Status**

Using only CLL-cell expression of ZAP-70 or IGHV mutation status, we could segregate patients into three subgroups that differed in their median times from diagnosis to initial therapy (Figure 3A). Patients in the validation cohort who had CLL cells that expressed ZAP-70 and used U-IGHV (N=307) had a median time from diagnosis to initial therapy of 2.5 years, which was not significantly different from that of 67 patients in the validation cohort who had ZAP-70-positive CLL cells that used M-IGHV (3.0 years) (Figure 3A). On the other hand, the 164 patients with ZAP-70-negative CLL cells that used U-IGHV (N=164, or 23% of the total) had a median time from diagnosis to initial treatment of 6.3 years, whereas the 167 patients with ZAP-70-negative CLL cells that expressed M-IGHV (N=167, or 24% of the total) had a highly indolent clinical course with a median time from diagnosis to first treatment of 10 years. These three subgroups also could be delineated among the subset of 521 patients in the validation cohort who contributed samples prior to initial treatment (Figure 3B). The patients in this subgroup who had CLL cells that expressed ZAP-70 and used U-IGHV (N=198) had a median time from diagnosis to initial therapy of 3.9 years, which was not significantly different from that of 49 patients in this subgroup who had ZAP-70-positive CLL cells that used M-IGHV (4.1 years) (Figure 3B). On the other hand, the 274 patients with ZAP-70-negative CLL cells could be segregated into a subgroup with CLL
cells that used U-IGHV that had a median time from diagnosis to initial treatment of 7.0 years, which was significantly longer than that of patients with ZAP-70-positive CLL cells, but significantly shorter than that of patients with ZAP-70-negative CLL cells that used M-IGHV (11.6 years). This hierarchy also was apparent for the subgroup of 199 patients with early-to-intermediate stage disease who had samples collected within one year of diagnosis (Figure 3C). The patients in this subgroup who had CLL cells that expressed ZAP-70 and used U-IGHV (N=77) had a median time from diagnosis to initial therapy of 3.1 years, which was not significantly different from that of 25 patients in this subgroup who had ZAP-70-positive CLL cells that used M-IGHV (3.2 years) (Figure 3C). On the other hand, the 97 patients with ZAP-70-negative CLL cells in this subgroup could be segregated into a subgroup of 52 patients with CLL cells that used U-IGHV who had a median time from diagnosis to initial treatment of 6.8 years, which was significantly longer than that of patients with ZAP-70-positive CLL cells, but shorter than that of patients with ZAP-70-negative CLL cells that used M-IGHV (time not determined (ND)) (Supplementary Table1).

**Relationship Between High-risk Genetic Features and CLL-cell Expression of ZAP-70 Or U-IGHV**

We examined the relationship between these prognostic markers and high-risk genetic features in the 323 patients in the validation cohort for whom we had cytogenetic data. Similar to the findings on the entire validation cohort, patients with CLL cells that expressed ZAP-70 (Figure 3D) had a significantly shorter time from diagnosis to initial therapy than did cases that lacked expression of ZAP-70, regardless of the IGHV
mutation status (p<0.001). Genomic aberrations were detected in 75% of these cases by interphase FISH (242/323). Ninety-four percent of the 48 CLL cell samples that were found to have high-risk genetic features of del17p13 or del11q22 also expressed ZAP-70 and used U-IGHV, whereas only 6% of such cases lacked expression of ZAP-70 (Fisher’s exact test, p<0.001) (Table 3). However, CLL cells that expressed ZAP-70, but used M-IGHV, did not have a higher frequency of del17p13 or del11q22, or higher use-frequency of IGHV3-21, than did ZAP-70-positive CLL cases that used U-IGHV, in contrast to previous studies29. Collectively, analyses of all these high-risk factors revealed that ZAP-70 was an independent predictor of the need for early treatment in CLL.
DISCUSSION

In this study we applied criteria for expression of ZAP-70, CD38, and IGHV mutation status defined in a test cohort to segregate patients in a validation cohort. Our validation cohort was comprised of CLL patients who were evaluated in a manner similar to that of our test cohort22. Thus, the analyses for ZAP-70, CD38, and IGHV mutation status were uniformly performed by the CRC for all 1,012 CLL cases examined. Among these three prognostic markers, we found that ZAP-70 was the strongest marker associated with a relatively short time interval from diagnosis to initial therapy. This was apparent even for the subgroup of 521 patients in the validation cohort who submitted blood samples to the CRC prior to receiving any treatment for CLL (Table 2, Figure 3B).

The median age at diagnosis of the patients who participated in our validation cohort was 56, which is substantially younger than the median age at diagnosis for 70 all patients with this disease28. The younger median age in our cohort reflects the demographics of patients who present to referral centers such as those in the CRC, as no selection for age was used in assigning patients to this study. To determine whether our findings also were relevant to older patients with this disease, we segregated out patients from our validation cohort who were 65 years or older at diagnosis who had a median age at diagnosis of 70 years. We found that the relative strength of each prognostic marker noted for the entire validation cohort also applied for this subgroup of patients with ages comparable to that of most patients with this disease (Table 2).

To test the predictive value of these markers in defining the time to initial therapy, we also segregated out early-to-intermediate stage patients (Rai stages 0-2) in
the validation cohort who submitted samples to the CRC within one year of diagnosis prior to receiving therapy for CLL (N = 199). For these 199 patients we found that CLL-cell expression of ZAP-70 still maintained its relative predictive value when compared with IGHV mutation status or CLL-cell expression of CD38 (Table 2, Figure 3C). We observed a significant difference in the number of patients with ZAP-70 positive CLL cells who eventually received treatment versus the number of patients with ZAP-70 negative CLL cells who did not receive treatment. This analysis revealed that CLL-cell expression of ZAP-70 was the strongest risk factor of the three for predicting a relatively short time from diagnosis to future therapy in patients with newly diagnosed disease.

Prior studies have found that ZAP-70 has functional significance in CLL\(^{30}\). Expression of ZAP-70 can enhance the intracellular signaling capacity of the Ig expressed in CLL\(^{31,32}\), a function that apparently is independent of its kinase activity\(^{33,34}\). This suggests a model proposing that CLL cells that express ZAP-70 are more responsive to self and/or environmental antigens that interact with the highly selected repertoire of Ig used by CLL B cells. Provided that there is repeated exposure to such antigens over time, this could allow for leukemia B cells stimulation leading to enhanced proliferation and/or resistance to apoptosis\(^{25,30}\). If so, then the expression of ZAP-70 might be more closely tied to the propensity for early disease progression than the mutation status of the expressed IGHV.

Some studies have not found ZAP-70 to be an independent risk factor for aggressive disease\(^{35-39}\). Conceivably, this might relate to the methods used for determining whether the CLL cells express this intracellular tyrosine kinase. Some methods, such as immunohistochemistry or immunoblot analysis\(^{23,40-44}\), might be less
quantitative than those using flow cytometry. On the other hand, flow cytometric analyses require that the CLL cells be made permeable to the antibodies used to detect ZAP-70. Assays using indirect immunofluorescence with non-conjugated anti-ZAP-70 antibodies might be subject to more variation than assays using directly labeled mAbs, further compounding this problem. Moreover, one recent study observed differences in the percent of leukemia cells staining positive for ZAP-70 when different mAbs were used on the same leukemia-cell population, even when the same gating strategy was used in the analysis. As such, analyses using other antibodies for ZAP-70 should be compared with those using the directly labeled IE7.2, as different anti-ZAP-70 mAb apparently can yield different results. That such standardization can be achieved was demonstrated in a recent multicenter study, indicating that optimized flow cytometry methods can yield reproducible and consistent results on leukemia-cell expression of ZAP-70 between different sites.

From the current study it appears that most patients with leukemia cells that express U-IGHV and/or CD38, but that lack expression of ZAP-70, will not require therapy by current criteria for many years after diagnosis. Moreover, approximately one-fourth of the patients who have CLL cells that use U-IGHV apparently lack leukemia cell-expression of ZAP-70. Such patients have a relatively indolent clinical course and might not enjoy the same risk-benefit ratio with early therapy as patients with CLL cells that express ZAP-70, who on average require therapy within 3 years of diagnosis. The latter high-risk group includes the approximately 10% of patients who have CLL cells that use M-IGHV.
Clinical studies are in progress to evaluate the potential benefit of early therapy in newly diagnosed patients who have adverse prognostic markers. Such studies might determine whether patients at high-risk for early disease progression benefit from therapy administered soon after diagnosis. Until the outcome of these studies are known, patients should not be treated on the basis of prognostic markers outside the context of a clinical trial. In addition, the use of these prognostic markers should not obviate clinical monitoring for other features associated with disease progression, such as lymphocyte doubling time, progressive lymphadenopathy, measurement of beta-2-microglobulin levels, or development of disease-related symptoms or disease-associated cytopenias. Also, because patients with early-stage CLL often do not have disease-related symptoms and/or are incidentally diagnosed, there is variability in the interval between the actual disease onset and diagnosis. Finally, some patients with leukemia cells that have high-risk features might have an exceptionally indolent course and visa versa. Nevertheless, with these caveats in mind, it appears appropriate to closely monitor those patients who have leukemia cells that express ZAP-70 for signs of disease progression, as they appear to be at high risk for requiring therapy within approximately 3 years of diagnosis as per current treatment guidelines.
ACKNOWLEDGMENTS

We are indebted to Lang Huynh, Traci Toy, and Mary Nguyen for their excellent technical assistance and to Monica Cook for data management. This work was supported by a grant (PO1-CA81534) from the National Institutes of Health to the Chronic Lymphocytic Leukemia Research Consortium. Dr. J.C. Byrd is a clinical scholar of the Leukemia and Lymphoma Society of America.

AUTHOR’S CONTRIBUTIONS

L.Z.R. designed and performed the research, analyzed the data, and wrote the manuscript. S.J., D.S.N., and F.H. analyzed the data, performed statistical analysis, and wrote the manuscript. W.G.W., M.J.K., J.G.G., J.R.B., K.R.R., J.C.B., M.R.G., N.E.K., contributed patient samples, data, and edited the manuscript. A.W.G. performed the database and bioinformatics analysis. T.J.K. designed the research, contributed patient samples, and wrote the manuscript. Authors declare no competing financial interests.
REFERENCES


Table 1

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Test cohort (N=307)</th>
<th>Validation cohort (N=705)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Median time (yrs) Dx to Tx</td>
</tr>
<tr>
<td>ZAP-70 negative</td>
<td>166 (54%)</td>
<td>9.1</td>
</tr>
<tr>
<td>ZAP-70 positive</td>
<td>141 (46%)</td>
<td>2.9</td>
</tr>
<tr>
<td>M-IGHV</td>
<td>143 (47%)</td>
<td>9.2</td>
</tr>
<tr>
<td>U-IGHV</td>
<td>164 (53%)</td>
<td>3.4</td>
</tr>
<tr>
<td>CD38 negative</td>
<td>207 (67%)</td>
<td>7.8</td>
</tr>
<tr>
<td>CD38 positive</td>
<td>100 (33%)</td>
<td>3.4</td>
</tr>
</tbody>
</table>
## Table 2

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Validation cohort (N=705)</th>
<th>Pre-treatment samples (N=521)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All (N=705)</td>
<td>Rai 0 (N=155)</td>
</tr>
<tr>
<td>ZAP-70</td>
<td>4.08*</td>
<td>4.45*</td>
</tr>
<tr>
<td>ZAP-70 /IGHV</td>
<td>3.09*</td>
<td>3.92*</td>
</tr>
<tr>
<td>ZAP-70/CD38</td>
<td>3.96*</td>
<td>4.46*</td>
</tr>
<tr>
<td>ZAP-70/IGHV/CD38</td>
<td>3.01*</td>
<td>3.97*</td>
</tr>
<tr>
<td>IGHV</td>
<td>3.14*</td>
<td>2.22†</td>
</tr>
<tr>
<td>IGHV/ZAP-70</td>
<td>1.87*</td>
<td>1.42</td>
</tr>
<tr>
<td>IGHV/CD38</td>
<td>3.05*</td>
<td>2.14†</td>
</tr>
<tr>
<td>IGHV/ZAP-70/CD38</td>
<td>1.86*</td>
<td>1.40</td>
</tr>
<tr>
<td>CD38</td>
<td>1.53*</td>
<td>0.57</td>
</tr>
<tr>
<td>CD38/ZAP-70</td>
<td>1.14</td>
<td>0.56</td>
</tr>
<tr>
<td>CD38/IGHV</td>
<td>1.39†</td>
<td>0.60</td>
</tr>
<tr>
<td>CD38/ZAP-70/IGHV</td>
<td>1.11</td>
<td>0.56</td>
</tr>
</tbody>
</table>

* p < 0.001
† p < 0.05
§ samples collected from patients within 1 yr of diagnosis
Table 3

<table>
<thead>
<tr>
<th>Genetic Characterization</th>
<th>Concordant cases</th>
<th></th>
<th>Discordant cases</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total # pts</td>
<td># pts</td>
<td># pts</td>
<td>P value</td>
</tr>
<tr>
<td>17p13q or 11q22</td>
<td>48</td>
<td>3 (6%)</td>
<td>45 (94%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>IGHV3-21</td>
<td>4</td>
<td>3 (75%)</td>
<td>1 (25%)</td>
<td>0.154</td>
</tr>
<tr>
<td>Other</td>
<td>167</td>
<td>77 (46%)</td>
<td>90 (54%)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>
TABLE 1

Biological And Clinical Characteristics Of The Patients

This table provides the data or number and percent in parentheses (where appropriate) of patients with characteristics listed in the left column for the test cohort (middle column) or validation cohort (right column).

TABLE 2

Comparison Of The Prognostic Significance Of ZAP-70, CD38, And IGHV Mutation Status On The Time To Initial Treatment In The Validation Cohort By Cox Proportional Hazards Univariate And Multivariable Regression Analyses

This table lists the HRs for cases with CLL cells that were ZAP-70 positive, expressed U-IGHV, or were positive for CD38 in a univariate analysis. This table also lists the HRs when these prognostic markers were combined in a multivariable analysis when one or both of the other prognostic markers were added to the model (i.e. ZAP-70 plus IGHV). The validation cohort is segregated into subgroups by Rai stage, by age at the time of sample collection, by the patients who were or were not treated for CLL at the time of sample collection, or by patients with Rai stage 0-2 who had samples collected within 1 year of diagnosis (median 0.24 years after diagnosis). The univariate and multivariable analyses were analyzed via the Cox Proportional Hazards Regression Model. Hazard Ratio (HR) is the ratio of the hazard functions between two groups. Here, the reference group in the HR is the set of patients having the adverse marker and an increase risk for early treatment. An asterisk (*) is used to denote HR that have
a p value of less than 0.001, a ( ) is used to denote a HR that has a p value of less than 0.05. The symbol () is used to denote the subgroup of patients with Rai stage 0-2 stage disease who had blood samples collected for these analyses within one year of diagnosis (median = 0.24) prior to receiving treatment.

**TABLE 3**

*Relationship Between ZAP-70, IGHV Mutation Status, And The Time From Diagnosis To Initial Therapy In The Subset Of Patients In The Validation Cohort With Cytogenetic Data (N=323)*

The distribution of cases according to the ZAP-70 expression and IGHV mutation status (concordant and discordant) and according to the presence or absence of additional genetic high-risk features, such as del11q22, del17p13, or use of *IGHV3-21*, are listed.

- **ZAP-70**: zeta-associated protein of 70 kD
- **M-IGHV**: mutated IGHV
- **U-IGHV**: unmutated IGHV
- **No. pts**: number of patients in each group
- **P***: p values determined according to Fisher’s exact test
FIGURES LEGENDS

Figure 1

Relationship Between ZAP-70 And CD38 Or IGHV Mutation Status In Defining The Time From Diagnosis To Initial Therapy

Kaplan-Meier curves depict the proportion of untreated CLL patients in the validation cohort (N=705) from diagnosis to initial therapy. Panels A, B, C, and D depict the time from diagnosis to initial treatment of patients who have CLL cells that express ZAP-70 (ZAP-70⁺), whereas panels E, F, G, and H show the time from diagnosis to first treatment of patients with CLL cells that were ZAP-70-negative (ZAP-70⁻). Panels A and E or B and F show how IGHV mutation status can segregate cases that are CD38-positive (CD38⁺) or CD38-negative (CD38⁻), respectively. Panels C and G or D and H show how CD38 can segregate cases that use U-IGHV or M-IGHV, respectively. The p values were determined using the log-rank test. The symbols represent the time at which patients were censored.

Figure 2

Relationship Between IGHV Mutation Status And ZAP-70 Or CD38 In Defining The Time From Diagnosis To Initial Therapy

Kaplan-Meier curves depict the proportion of untreated CLL patients in the validation cohort (N = 705) from diagnosis to initial therapy. Panels A, B, C, and D depict the time from diagnosis to initial treatment of patients who have CLL cells that express M-IGHV genes, whereas panels E, F, G, and H show the time from diagnosis to first treatment of patients with CLL cells use U-IGHV. Panels A and E or B and F show how
CD38 can segregate cases that are ZAP-70-negative (ZAP-70⁻) or ZAP-70-positive (ZAP-70⁺), respectively. Panels C and G or D and H show how ZAP-70 can segregate cases that are CD38-positive (CD38⁺) or CD38-negative (CD38⁻), respectively. The p values were determined using the log-rank test. The symbols represent the time at which patients were censored.

**Figure 3**

*Relationship Between ZAP-70 and IGHV Mutation Status In Defining The Time From Diagnosis To Initial Therapy*

Kaplan-Meier curves depict the proportion of untreated patients with CLL according to the time since diagnosis. The symbols represent the time at which patients were censored. The patients were divided into 4 subgroups according to their ZAP-70 expression and IGHV mutation status. The legend is provided in each panel for these 4 subgroups, namely cases that were ZAP-70-negative and that used M-IGHV (ZAP-70⁻/M-VH) with closed boxes for symbols, cases that were ZAP-70-negative but used U-IGHV (ZAP-70⁻/U-VH) with closed triangles for symbols, cases that were ZAP-70-positive but used M-IGHV (ZAP-70⁺/M-VH) with inverted triangles for symbols, and cases that were ZAP-70-positive and used U-IGHV (ZAP-70⁺/U-VH) with closed diamonds for symbols. The median time from diagnosis to initial therapy for subgroup is listed in parenthesis in the legend. The number of patients in each subgroup is provided in the legend. **Panel A** provides data for the entire validation cohort of 705 patients. **Panel B** provide data for the 521 patients in the validation cohort who contributed samples for analyses prior to receiving therapy for CLL. **Panel C** provides
data for the 199 patients who presented with early-to-intermediate stage disease within one year of diagnosis prior to receiving therapy for CLL. **Panel D** provides data for the 323 patients who had cytogenetic analyses performed on their leukemia cells at the time of sample collection (Supplementary Table 1).
Figure 1
Figure 2
Figure 3
Relative value of ZAP-70, CD38, and immunoglobulin mutation status in predicting aggressive disease in chronic lymphocytic leukemia