THE USE OF MOLECULAR PROFILING TO PREDICT SURVIVAL AFTER CHEMOTHERAPY FOR DIFFUSE LARGE-B-CELL LYMPHOMA

ANDREAS ROSENWALD, M.D., GEORGE WRIGHT, PH.D., WING C. CHAN, M.D., JOSEPH M. CONNORS, M.D., ELIAS CAMPO, M.D., RICHARD I. FISHER, M.D., RANDY D. GASCOYNE, M.D., H. KONRAD MULLER-HERMELINK, M.D., ERLEND B. SMELAND, M.D., PH.D., AND LOUIS M. STAUDT, M.D., PH.D., FOR THE LYMPHOMA/LEUKEMIA MOLECULAR PROFILING PROJECT

ABSTRACT

Background The survival of patients with diffuse large-B-cell lymphoma after chemotherapy is influenced by molecular features of the tumors. We used the gene-expression profiles of these lymphomas to develop a molecular predictor of survival.

Methods Biopsy samples of diffuse large-B-cell lymphoma from 240 patients were examined for gene expression with the use of DNA microarrays and analyzed for genomic abnormalities. Subgroups with distinct gene-expression profiles were defined on the basis of hierarchical clustering. A molecular predictor of risk was constructed with the use of genes with expression patterns that were associated with survival in a preliminary group of 160 patients and was then tested in a validation group of 80 patients. The accuracy of this predictor was compared with that of the international prognostic index.

Results Three gene-expression subgroups — germinal-center B-cell–like, activated B-cell–like, and type 3 diffuse large-B-cell lymphoma — were identified. Two common oncogenic events in diffuse large-B-cell lymphoma, bcl-2 translocation and c-rel amplification, were detected only in the germinal-center B-cell–like subgroup. Patients in this subgroup had the highest five-year survival rate. To identify other molecular determinants of outcome, we searched for individual genes with expression patterns that correlated with survival in the preliminary group of patients. Most of these genes fell within four gene-expression signatures characteristic of germinal-center B cells, proliferating cells, reactive stromal and immune cells in the lymph node, or major-histocompatibility-complex class II complex. We used 17 genes to construct a predictor of overall survival after chemotherapy. This gene-based predictor and the international prognostic index were independent prognostic indicators.

Conclusions DNA microarrays can be used to formulate a molecular predictor of survival after chemotherapy for diffuse large-B-cell lymphoma. (N Engl J Med 2002;346:1937-47.)

Copyright © 2002 Massachusetts Medical Society.
Hierarchical clustering was used to define subgroups of diffuse large-B-cell lymphoma. To create an outcome variable based on gene-expression studies, we divided the patients into two groups: the preliminary group comprised 160 patients, and the validation group comprised 80 patients. Within the preliminary group, the significance of the correlation between outcome (overall survival after chemotherapy) and gene-expression data from individual microarray features was determined with use of the Cox proportional-hazards model. Genes that were associated with a good or a bad outcome at a significance level of P<0.01 were assigned to gene-expression signatures, as described previously. These representative genes were chosen because of their high variance in expression (i.e., they were among the top third of genes in variance of gene-expression levels in the preliminary group). We used the average value for each signature and the value for the BMP6 gene, a member of the transforming growth factor β superfamily of genes, to construct a multivariate Cox survival model according to the following formula: outcome-predictor score = (0.241 × the average value of the proliferation signature) + (0.310 × the average value of the germinal-center B-cell signature) − (0.311 × the average value of the major-histocompatibility-complex [MHC] class II signature) − (0.249 × the average value of the lymph-node signature). The coefficients in this formula were derived from the Cox model, and a high score indicates a poor outcome. We used the Wald test to determine the significance of the association between this model and the outcome in the preliminary group, the validation group, and the group as a whole and to assess the independence of the risk groups defined by the international prognostic index and the outcome predicted on gene-expression profiles. All t-tests were two-sided except those used for the validation group; the results of one-sided t-tests are reported for this group, since the analysis of the preliminary group indicated the direction of the effect. The data set used to determine the outcome predictor and a detailed description of the statistical methods used are available as Supplementary Appendix I with the full text of this article at http://www.nejm.org and at http://llmpp.nih.gov/DLBCL.

RESULTS

Molecular, Pathological, and Clinical Features of the Subgroups

We first determined whether we could identify the previously described gene-expression subgroups in the group of diffuse large-B-cell lymphomas that we analyzed. We applied a hierarchical-clustering algorithm to group the lymphomas according to the expression of 100 genes that distinguished between germinal-center B-cell–like and activated B-cell–like diffuse large-B-cell lymphomas at a significance level of P<0.001 in the previous analysis and found three...
MOLECULAR PROFILING OF SURVIVAL IN DIFFUSE LARGE-B-CELL LYMPHOMA

A

Gene

A-myb  LMO2
Jnk3  CD10
bcl-6  Cyclin D2
IRF-4  Flip  CD44

Germinal-center B-cell–like  Type 3  Activated B-cell–like

Subgroup of Diffuse Large-B-Cell Lymphoma

Relative Level of Expression (× median value)

0.25  0.50  1.00  2.00  4.00

B

Oncogenic Abnormality

<table>
<thead>
<tr>
<th>Genomic Abnormality</th>
<th>Germinal-center B-cell–like</th>
<th>Type 3</th>
<th>Activated B-cell–like</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>no. of samples</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c-rel amplification</td>
<td>17</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>bcl-2 t(14;18)</td>
<td>26</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

C

Probability of Survival

P<0.001

NO. AT RISK

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>0</th>
<th>2</th>
<th>4</th>
<th>6</th>
<th>8</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Germinal-center B-cell–like</td>
<td>115</td>
<td>81</td>
<td>60</td>
<td>46</td>
<td>32</td>
<td>19</td>
</tr>
<tr>
<td>Type 3</td>
<td>52</td>
<td>24</td>
<td>18</td>
<td>10</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>Activated B-cell–like</td>
<td>73</td>
<td>35</td>
<td>23</td>
<td>19</td>
<td>8</td>
<td>5</td>
</tr>
</tbody>
</table>
large subgroups (Fig. 1A). One had a high level of expression of the genes characteristic of germinal-center B-cell–like diffuse large-B-cell lymphoma and normal germinal-center B cells, another expressed genes characteristic of activated B-cell–like diffuse large-B-cell lymphoma and mitogenically activated blood B cells, and the third, termed type 3 diffuse large-B-cell lymphoma, did not express either set of genes at a high level. The heterogeneity of expression within this third subgroup indicates that it may consist of more than one type of diffuse large-B-cell lymphoma.

The subgroups differed substantially with respect to two recurrent oncogenic events. The t(14;18) translocation involving the \textit{bcl-2} gene and the amplification of the \textit{c-rel} locus on chromosome 2p occurred exclusively in germinal-center B-cell–like diffuse large-B-cell lymphomas (Fig. 1B). These findings support the view that the various subgroups represent different diseases that arise as a result of distinct mechanisms of malignant transformation.\textsuperscript{3,11}

The clinical and pathological features of the three subgroups are shown in Table 1. The most common histologic type of diffuse large-B-cell lymphoma — centroblastic monomorphic — was found in 66 percent of the germinal-center B-cell–like subgroup but also in 32 percent of the activated B-cell–like subgroup and 27 percent of the type 3 subgroup. Centroblastic polymorphic and immunoblastic subtypes were more common in activated B-cell–like and type 3 diffuse large-B-cell lymphomas, but they were also observed in the germinal-center B-cell–like subgroup. Thus, these three subgroups were not strictly related to a specific histologic subtype. With respect to clinical features, a significantly higher proportion of patients in the activated B-cell–like subgroup than in the other two groups were older than 60 years of age (\(P=0.05\)) and had an ECOG performance status of more than 1 (\(P=0.03\)), but the tumor subgroup did not correlate with the risk groups defined on the basis of the international prognostic index (\(P=0.44\)).

Overall survival after anthracycline-based chemotherapy differed significantly among the three subgroups (\(P<0.001\)) (Fig. 1C). Patients with germinal-center B-cell–like diffuse large-B-cell lymphoma had a five-year survival rate of 60 percent, as compared with a rate of 39 percent for patients with diffuse type 3 large-B-cell lymphoma and 35 percent for those with activated B-cell–like diffuse large-B-cell lymphoma.

**Correlation between Expression of Individual Genes and Outcome**

Although the three subgroups may be viewed as distinct diseases, these divisions did not fully reflect the differences in survival among patients with diffuse large-B-cell lymphoma. For example, patients with germinal-center B-cell–like diffuse large-B-cell lymphoma

---

**Table 1. Characteristics of Patients with Diffuse Large-B-Cell Lymphoma.**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Total (N=240)</th>
<th>Germinal-center B-cell–like (N=115)</th>
<th>Type 3 B-cell–like (N=52)</th>
<th>Activated B-cell–like (N=73)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>International-prognostic-index component</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lactate dehydrogenase &gt;1× normal</td>
<td>57</td>
<td>56</td>
<td>50</td>
<td>65</td>
<td>0.3</td>
</tr>
<tr>
<td>Age &gt;60 yr</td>
<td>55</td>
<td>48</td>
<td>58</td>
<td>66</td>
<td>0.05</td>
</tr>
<tr>
<td>Ann Arbor stage &gt;II</td>
<td>55</td>
<td>53</td>
<td>53</td>
<td>59</td>
<td>0.68</td>
</tr>
<tr>
<td>No. of extranodal sites &gt;1</td>
<td>20</td>
<td>20</td>
<td>25</td>
<td>15</td>
<td>0.34</td>
</tr>
<tr>
<td>ECOG performance status &gt;1</td>
<td>22</td>
<td>19</td>
<td>16</td>
<td>33</td>
<td>0.03</td>
</tr>
<tr>
<td>Risk group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.44</td>
</tr>
<tr>
<td>Low (score, 0–1)</td>
<td>37</td>
<td>40</td>
<td>41</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>Intermediate (score, 2–3)</td>
<td>49</td>
<td>49</td>
<td>43</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>High (score, 4–5)</td>
<td>14</td>
<td>11</td>
<td>15</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Histologic subtype</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Centroblastic monomorphic</td>
<td>47</td>
<td>66</td>
<td>27</td>
<td>32</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Centroblastic polymorphic</td>
<td>19</td>
<td>9</td>
<td>27</td>
<td>28</td>
<td>0.009</td>
</tr>
<tr>
<td>Immunoblastic</td>
<td>8</td>
<td>3</td>
<td>12</td>
<td>15</td>
<td>0.007</td>
</tr>
<tr>
<td>Burkitt-like</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>8</td>
<td>0.31</td>
</tr>
<tr>
<td>Plasmablastic</td>
<td>5</td>
<td>5</td>
<td>8</td>
<td>6</td>
<td>0.71</td>
</tr>
<tr>
<td>T-cell rich</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0.05</td>
</tr>
<tr>
<td>Anaplastic</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0.05</td>
</tr>
<tr>
<td>Unclassified</td>
<td>15</td>
<td>18</td>
<td>16</td>
<td>11</td>
<td>0.49</td>
</tr>
</tbody>
</table>

*Because of rounding, percentages may not total 100. ECOG denotes Eastern Cooperative Oncology Group.
**TABLE 2. USE OF GENE-EXPRESSION PROFILES TO PREDICT OUTCOME FOR PATIENTS WITH DIFFUSE LARGE-B-CELL LYMPHOMA.**

<table>
<thead>
<tr>
<th>GENE-EXPRESSION VARIABLE</th>
<th>NO. OF MICROARRAY FEATURES IN SIGNATURE (N=7399)</th>
<th>NO. OF MICROARRAY FEATURES SIGNIFICANTLY ASSOCIATED (P&lt;0.01) WITH A GOOD OUTCOME (N=162)</th>
<th>NO. OF MICROARRAY FEATURES SIGNIFICANTLY ASSOCIATED (P&lt;0.01) WITH A BAD OUTCOME (N=508)</th>
<th>REPRESENTATIVE GENES</th>
<th>GENBANK ACCESSION NO.</th>
<th>P VALUE PRELIMINARY GROUP</th>
<th>P VALUE VALIDATION GROUP</th>
<th>P VALUE ALL PATIENTS (95% CI)</th>
<th>OUTCOME</th>
</tr>
</thead>
<tbody>
<tr>
<td>Germinal-center B-cell signature</td>
<td>151</td>
<td>15</td>
<td>0</td>
<td><em>bcl-6</em></td>
<td>IMAGE 1334260† IMAGE 814622†</td>
<td>U00115</td>
<td>AA805575</td>
<td>0.009</td>
<td>0.001</td>
</tr>
<tr>
<td>MHC class II signature</td>
<td>37</td>
<td>35</td>
<td>0</td>
<td>HLA-DP</td>
<td>HLA-DQ</td>
<td>HLA-DRα</td>
<td>HLA-DRβ</td>
<td>X00457</td>
<td>X00452</td>
</tr>
<tr>
<td>Lymph-node signature</td>
<td>357</td>
<td>30</td>
<td>2</td>
<td>α-Actinin</td>
<td>Collagen type III α 1</td>
<td>Connective-tissue growth factor</td>
<td>Fibronectin</td>
<td>KIAA0233</td>
<td>Urokinase plasminogen activator</td>
</tr>
<tr>
<td>Proliferation signature</td>
<td>1333</td>
<td>6</td>
<td>287</td>
<td><em>c-myc</em></td>
<td>E2F1G3</td>
<td>NPM</td>
<td></td>
<td>V00568</td>
<td>NM_014366</td>
</tr>
<tr>
<td>Other</td>
<td>5521</td>
<td>76</td>
<td>219</td>
<td>BMP6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*P values indicate the significance of the association between the average value of the gene-expression variable and overall survival. Relative risk indicates the change in risk associated with a change by a factor of 2 in the average value for a given gene-expression signature. CI denotes confidence interval, and MHC major histocompatibility complex.

†IMAGE refers to an Integrated Molecular Analysis of Genomes and Their Expression consortium complementary DNA clone on the Lymphochip microarray.
had the best prognosis, but these patients still had a 36 percent risk of death within three years after treatment. The patients with activated B-cell–like diffuse large-B-cell lymphoma, by contrast, had the worst prognosis, but the five-year survival rate of 35 percent suggests that some patients in this subgroup may be cured by chemotherapy.

For these reasons, we used a Cox proportional-hazards model to identify individual genes whose expression correlated with the outcome. Data from 670 of 7399 microarray features were significantly associated with a good or a bad outcome in the preliminary group (P<0.01); this number is greater than would be expected by chance (P=0.005 with the use of a permutation test) (Table 2).

To classify the genes that were correlated with outcome, we used hierarchical clustering to group them into gene-expression signatures. A gene-expression signature is a group of genes expressed in a specific cell lineage or stage of differentiation or during a particular biologic response. Many of the genes we identified fell within previously described gene-expression signatures (Table 2) (see Supplementary Appendix 1 at http://www.nejm.org). Among the 162 microarray features associated with a favorable outcome, 15 belonged to the signature that characterizes normal germinal-center B cells,3 30 were in the lymph-node signature of reactive nonmalignant cells in biopsy specimens of diffuse large-B-cell lymphoma,8 and 35 were in the MHC class II signature. In the proliferation signature, which includes genes that are highly expressed in dividing cells,4 287 of 1333 microarray features were associated with a poor outcome.

Since genes within the same gene-expression signature are probably associated with similar biologic aspects of a tumor, we combined the genes that were significantly associated with survival (P<0.01) within each signature. To minimize the number of genes in the outcome predictor, we selected 16 genes that were highly variable in expression — 3 germinal-center B-cell genes, 4 MHC class II genes, 6 lymph-node genes, and 3 proliferation genes — and averaged the expression values for genes belonging to the same signature (Table 2). In a univariate analysis, these four
MOLECULAR PROFILING OF SURVIVAL IN DIFFUSE LARGE-B-CELL LYMPHOMA

A

All Patients

Probability of Survival

Quartiles 1 and 2
Quartiles 3 and 4

IPI score, 0–1
IPI score, 2–3
IPI score, 4–5

P<0.001

Years

No. at Risk
IPI score, 0–1 82 62 50 37 22 13
IPI score, 2–3 108 56 38 28 19 13
IPI score, 4–5 32 10 6 4 3 1

B

Preliminary Group

Validation Group

All Patients

Probability of Survival

Quartiles 1 and 2
Quartiles 3 and 4

IPI Score, 0–1

P<0.001

Years

No. at Risk
Quartiles 1 and 2 38 31 28 20 11 7
Quartiles 3 and 4 17 10 7 5 3 2

IPI Score, 2–3

Quartiles 1 and 2
Quartiles 3 and 4

P<0.001

Years

No. at Risk
Quartiles 1 and 2 30 24 15 11 6 3
Quartiles 3 and 4 42 16 9 8 6 4

IPI Score, 4–5

Quartiles 1 and 2
Quartiles 3 and 4

P=0.06

Years

No. at Risk
Quartiles 1 and 2 8 4 2 1 0 0
Quartiles 3 and 4 14 3 1 0 0 0

P=0.13

P=0.02
signatures were found to predict survival in the preliminary group, in the validation group, and in the group of all patients (Table 2).

Individual genes that were not in these four signatures but that predicted the likelihood of survival in a univariate analysis of the preliminary group (P<0.01) were evaluated to determine whether these variables increased the predictive value of the test. The only gene that significantly increased the predictive value (P=0.005) was BMP6, which was associated with a poor outcome.

The final model combined the four gene-expression signatures and BMP6. Each case of diffuse large-B-cell lymphoma was assigned a score that was calculated as the weighted sum of these components, optimized by a Cox proportional-hazards model of overall survival within the preliminary group. The score, expressed as a continuous variable, correlated significantly with the clinical outcome in both the preliminary group (P<0.001) and the validation group (P<0.001), indicating that the results are reproducible. The score ranged from –1.7 to 2.4, with a standard deviation of 0.72, and each unit increase in the score induced an increase in the relative risk of death by a factor of 2.7 (95 percent confidence interval, 2.11 to 3.51).

The patients were ranked according to their score and divided into quartiles (from highest to lowest scores). Kaplan-Meier plots of overall survival showed distinct differences in the five-year survival rates in the various quartiles in both the preliminary and validation groups (Fig. 2). In the group as a whole, the five-year survival rates were 73 percent in quartile 1, 71 percent in quartile 2, 34 percent in quartile 3, and 15 percent in quartile 4 (Fig. 2C).

Comparison of the Gene-Expression–Based Outcome Predictor and the International Prognostic Index

The international-prognostic-index scores (Fig. 3A) and the gene-expression–based scores were independent predictors of survival in the preliminary group (P<0.001) and the validation group (P=0.002). In a multivariate Cox model that combined both the international-prognostic-index scores and the gene-expression–based scores, each unit increase in the latter score increased the relative risk of death by a factor of 2.6 (95 percent confidence interval, 2.02 to 3.48).

Kaplan-Meier plots of overall survival showed the independence of the international-prognostic-index score and the gene-expression–based score (Fig. 3B). For these plots, we combined quartiles 1 and 2 into one group and quartiles 3 and 4 into a second group. These two groups had significantly different outcomes in the analysis of patients with low or intermediate risk according to their international-prognostic-index scores, and this difference was observed in both the preliminary and the validation groups. The gene-expression–based method also identified the few patients in the high-risk group according to the international-prognostic-index score who were long-term survivors (Fig. 3B).

Relation between Gene-Expression–Based Score and Subtype of Diffuse Large-B-Cell Lymphoma

Since overall survival differed in the three subgroups of diffuse large-B-cell lymphoma (Fig. 1C), we investigated whether the components of the outcome predictor were differentially expressed by these subgroups. As expected, the germinal-center B-cell signature was much more highly expressed in the germinal-center B-cell-like subgroup than in the other two subgroups (Fig. 4A). The activated B-cell–like subgroup had the highest level of expression of the proliferation signature and BMP6 but the lowest level of expression of the lymph-node signature. On the other hand, the level of expression of the MHC class II signature was similar among the three subgroups. The gene-expression–based score was highest in the activated B-cell–like group, intermediate in the type 3 subgroup, and lowest in the germinal-center B-cell–like subgroup (Fig. 4A), demonstrating that this approach incorporates the clinical distinctions among the subgroups of diffuse large-B-cell lymphoma.

Nonetheless, the components of the outcome predictor were differentially expressed within each subgroup (Fig. 4B), and the predictor score could be used to subdivide the patients within each subgroup into distinct risk groups (Fig. 4C). This feature accounts for the predictor’s greater prognostic power as compared with that derived from the use of subgroups of diffuse large-B-cell lymphoma.

DISCUSSION

We have developed a method to predict the likelihood of survival after chemotherapy for diffuse large-B-cell lymphoma that is based on patterns of gene expression in biopsy specimens of the lymphoma. This molecular method and the clinically based international prognostic index were found to be independent predictors of the prognosis. The international prognostic index has not proved effective in stratifying patients with diffuse large-B-cell lymphoma for therapeutic trials, but we think that the gene-expression–based method might serve this purpose.

Our analysis of gene-expression profiles integrated two complementary approaches to outcome prediction. In the first approach, we used hierarchical clustering to identify subgroups that differed with respect to the expression of hundreds of genes. With the use of this clustering method, two recurrent genetic abnormalities — the t(14;18) translocation and amplification of the c-rel locus — were found exclusively in the germinal-center B-cell–like subgroup. This finding sup-
Figure 4. Relation between the Gene-Expression–Based Outcome-Predictor Score and the Subgroup of Diffuse Large-B-Cell Lymphoma.

Panel A shows the mean (±SE) expression value (after log₂ transformation) of each outcome-predictor component and the score. Panel B shows the level of expression of variables in the outcome predictor and the scores in the three subgroups of lymphoma. Panel C shows Kaplan–Meier plots of overall survival in the subgroups, stratified according to the quartile of risk reflected by the gene-expression–based score. Quartiles 1 and 2 and quartiles 3 and 4 were merged. Higher quartiles indicate a poorer outcome. MHC denotes major histocompatibility complex.
ports our view that the subgroups of diffuse large-B-cell lymphoma represent distinct disease entities. Further support for this idea is supplied by the finding that the germinal-center B-cell–like subgroup had a significantly greater likelihood of survival after chemotherapy than did the activated B-cell–like and type 3 subgroups.

In the second approach, we used clinical and gene-expression data to identify individual genes that predicted the outcome and then combined these variables into a multivariate model. This model incorporated differences in the levels of gene expression among the subgroups of diffuse large-B-cell lymphoma that influenced the outcome, as well as other differences in gene expression that were associated with the likelihood of survival.

The predictive genes fell within four biologic groups defined on the basis of gene-expression signatures. The proliferation signature was the best predictor of an adverse outcome — a finding that is consistent with those of previous analyses of tumor-cell proliferation in diffuse large-B-cell lymphoma. Two of the gene-expression signatures that were associated with a good outcome suggest that the immune response to the tumor cells may be a crucial determinant of survival after chemotherapy. The MHC class II gene-expression signature correlated with a good outcome, suggesting that antigen presentation to the immune system has a role in therapeutic responses. The genes in the lymph-node signature that were associated with a good outcome encode components of the extracellular matrix and connective-tissue growth factor, a mediator of fibrosis that promotes the synthesis of the matrix. Sclerotic reactions occur in some cases of diffuse large-B-cell lymphoma, and the lymph-node signature may reflect these reactions. Other genes in the lymph-node signature are characteristically expressed in macrophages and natural killer cells, again suggesting that an antitumor immune response improves survival after chemotherapy.

The prognostically favorable germinal-center B-cell gene-expression signature is notable because cell lines derived from germinal-center B-cell–like diffuse large-B-cell lymphomas have decreased activity of the nuclear factor κB signaling pathway. This protective pathway interferes with the apoptotic effect of chemotherapy. In contrast, in activated B-cell–like diffuse large-B-cell lymphomas, there is constitutive activation of this pathway, which, in principle, could block the apoptosis induced by chemotherapy and thus account for the relatively poor outcome in this subgroup.

Our outcome predictor may help identify patients with diffuse large-B-cell lymphoma who are unlikely to be cured by conventional therapy. On the basis of their predictor scores, one quarter of the 240 patients in this study could be assigned to a risk group with a five-year survival rate of 15 percent. Among patients at intermediate risk according to the international-prognostic-index score, use of the outcome predictor indicated that 55 percent of these patients had a five-year survival rate of 18 percent, whereas this group of patients had a five-year survival rate of 41 percent overall. Even among patients with a low risk according to the international-prognostic-index score, use of the outcome predictor showed that 16 percent of the patients in this group had a five-year survival rate of only 28 percent (data not shown).

One of the features of our analysis is that the outcome predictor involves a small number of genes and thus multiplexed quantitative reverse-transcriptase–polymerase-chain-reaction assays or customized DNA microarrays could easily be developed for clinical application. Regardless of the eventual choice of technique, our study highlights the need to use molecular diagnosis in patients with diffuse large-B-cell lymphoma, since these patients have molecularly distinct diseases that may require individualized and molecularly targeted therapies.

Supported by a Director’s Challenge grant (UO1-CA84967) from the National Cancer Institute, by the Center for Cancer Research of the National Cancer Institute, and by the Cancer Genome Anatomy Project. Dr. Rosenwald was supported in part by the Deutsche Krebshilfe, Bonn, Germany. Ms. Averett was a Howard Hughes Medical Institute Research Scholar at the National Institutes of Health during the study.

We are indebted to Bob Strausberg, Narajan Bhat, and Butch Hopkins for preparing the messenger RNA reference pool and for helpful discussions, and to Ken Buettow and Jenny Kelly for verifying the sequence of Lymphochip clones.

REFERENCES


