Prognostic Biomarkers in Diffuse Large B-Cell Lymphoma

Izidore S. Lossos and Daniel Morgensztern

ABSTRACT

Diffuse large B-cell lymphoma (DLBCL) is the most common type of non-Hodgkin’s lymphoma. Although it represents a curable disease, less than half of the patients are cured with conventional chemotherapy. The highly variable outcome reflects a heterogeneous group of tumors, with different genetic abnormalities and response to therapy. The International Prognostic Index (IPI) is useful in predicting the outcome of DLBCL patients. However, patients with identical IPI still exhibit marked variability in survival, suggesting the presence of significant residual heterogeneity within each IPI category. The discovery of specific genetic alterations and the assessment of protein expression led to the identification of multiple novel single molecular markers capable of predicting the outcome of DLBCL patients independently of clinical variables. The recent application of DNA microarrays and tissue array technologies allowed a better understanding of the biology of lymphoma and the development of novel diagnostic tools capable of improving the current models for outcome prediction. However, much confusion exists in the literature regarding the importance of different prognostic biomarkers and their applicability in routine practice. This review summarizes the recent advances in our understanding of prognostic biomarkers in DLBCL and discusses whether this is the right time for biomarkers-guided risk-adjusted therapy.

J Clin Oncol 24:995-1007. © 2006 by American Society of Clinical Oncology

INTRODUCTION

Diffuse large B-cell lymphoma (DLBCL) is the most common type of lymphoma, with an annual incidence of more than 25,000 occurrences in the United States. It accounts for approximately one third of the total number of adult non-Hodgkin’s lymphoma (NHL) patients. Initially grouped into the intermediate prognostic grade by the International Working Formulation (IWF), DLBCL is recognized as a distinct entity by the Revised European-American Lymphoma and WHO classifications.

Patients with DLBCL have highly variable outcomes reflecting a heterogeneous group of tumors with different genetic abnormalities, clinical features, response to treatment, and prognosis. Combination chemotherapy containing anthracyclines has transformed DLBCL from a universally fatal disease to a potentially curable one. Although most patients respond to initial anthracycline-based therapy, fewer than half are cured. Identification of patients who do not benefit from current treatment may constitute the basis for risk-adjusted therapies for DLBCL. It could also lead to identification of patients who may be candidates for investigational approaches, and allow examination whether new therapeutic approaches improve the outcome of high-risk patients without confounding the study population with patients who benefit from the standard therapy.

The management of patients with DLBCL had been traditionally guided mainly by the Ann Arbor staging, which was originally developed for patients with Hodgkin’s disease and is based on the contiguous spreading to adjacent lymph nodes. This staging system, however, is less accurate in NHL, which is associated with a more unpredictable behavior including early systemic dissemination and involvement of extranodal sites. Many investigators attempted to improve the predictive outcome for intermediate-grade lymphomas (IWF), including DLBCL, by grouping patients based on prognostic models that relied on clinical and laboratory features, which remained independent predictors of overall survival in multivariate analysis (Table 1 summarizes some of the proposed models).

The International Non-Hodgkin’s Lymphoma Prognostic Factors Project established a clinical predictor for overall survival (OS) in IWF categories F, G, and H that contain also DLBCL. The proposed International Prognostic Index (IPI) was based on five independent prognostic factors including age, Ann Arbor tumor stage, serum lactate dehydrogenase, performance status, and number of extranodal sites. The subdivision of patients according to the number of prognostic factors into low risk (none or one factor), low-intermediate risk (two factors), high-intermediate risk (three factors), or high risk (four or five factors) with predicted 5-year OS values of 73%, 51%, 43%, and 26%, rapidly became the most widely used and accepted prognostic model for intermediate-grade lymphoma. The prognostic value of the IPI has been validated subsequently in patients with DLBCL.
In all the clinical models, including the IPI index, there was marked residual heterogeneity in outcome, as reflected by considerably variable survival of patients with identical prognostic scores. The latter was attributed to the marked genetic and molecular heterogeneity that underlies disease aggressiveness and tumor progression, and led to evaluation of molecular and genetic markers associated with patients’ survival.

Herein, we review the current knowledge of prognostic biomarkers in DLBCL. We initially summarize individual biomarkers identified in studies evaluating prognostic significance of specific genes or proteins. This is followed by the review of studies simultaneously evaluating the prognostic value of multiple biomarkers, thus accounting for possible biologic interactions between the individual prognostic biomarkers. Finally, we discuss whether this is the right time for incorporation of biomarkers-guided risk-adjusted therapy into clinical trials.

**Table 1. Sample of Clinical Prognostic Models in Intermediate-Grade Lymphoma**

<table>
<thead>
<tr>
<th>Author (affiliation, year)</th>
<th>No. of Patients</th>
<th>Risk Factors</th>
<th>5-Year Survival</th>
<th>No. of Risk Factors</th>
<th>% Survival</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shipp9 (Dana-Farber, 1986)</td>
<td>121</td>
<td>Performance status 0</td>
<td>68</td>
<td>Bulky disease 1 (except PS)</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bulky disease 1 (except PS)</td>
<td>55</td>
<td>Extranodal sites 1 to 3</td>
<td>24</td>
</tr>
<tr>
<td>Jagannath10 (M.D. Anderson, 1986)</td>
<td>105</td>
<td>LDH 0</td>
<td>87</td>
<td>Tumor burden* 1</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tumor burden* 1</td>
<td>48</td>
<td>2</td>
<td>20</td>
</tr>
<tr>
<td>Velasquez11 (M.D. Anderson, 1989)</td>
<td>250</td>
<td>Age 0</td>
<td>85</td>
<td>Tumor burden 1</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tumor burden 1</td>
<td>68</td>
<td>LDH 2</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LDH 2</td>
<td>39</td>
<td>Extranodal sites 3</td>
<td>18</td>
</tr>
<tr>
<td>Swan12 (M.D. Anderson, 1989)</td>
<td>86</td>
<td>LDH 0</td>
<td>100</td>
<td>β2-microglobulin 1</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td></td>
<td>β2-microglobulin 1</td>
<td>54</td>
<td>2</td>
<td>19</td>
</tr>
<tr>
<td>Coiffier14 (GELA, 1991)</td>
<td>737</td>
<td>LDH</td>
<td>88*</td>
<td>Stage* 0</td>
<td>71*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stage* 0</td>
<td>71*</td>
<td>Extranodal sites 1 or 2 (or LDH alone)</td>
<td>41*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tumor bulk 1 or 2 (or LDH alone)</td>
<td>41*</td>
<td>1 to 3 (with LDH)</td>
<td>52</td>
</tr>
<tr>
<td>Rodriguez15 (M.D. Anderson, 1992)</td>
<td>144</td>
<td>β2-microglobulin†</td>
<td>83†</td>
<td>LDH 0 to 2</td>
<td>24†</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stage 3 to 5</td>
<td>24†</td>
<td>Bulky disease</td>
<td>24†</td>
</tr>
<tr>
<td></td>
<td></td>
<td>“B” symptoms</td>
<td>24†</td>
<td></td>
<td>24†</td>
</tr>
<tr>
<td>Shipp127 (IPI, 1993)</td>
<td>2031</td>
<td>Age (≤ 60 v &gt; 60 years) 0 or 1</td>
<td>73</td>
<td>Stage I/II v III/IV</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stage I/II v III/IV 2</td>
<td>51</td>
<td>PS (0/1 v 2-4) 3</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PS (0/1 v 2-4) 3</td>
<td>43</td>
<td>LDH (normal v increased) 4 or 5</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Extrapolated sites (≠ 1 v &gt; 1) 4 or 5</td>
<td>26</td>
<td></td>
<td>26</td>
</tr>
<tr>
<td>Conconi16 (IPI-β2- microglobulin, 2000)</td>
<td>111</td>
<td>IPI, β2-microglobulin 0 or 1</td>
<td>NA†</td>
<td></td>
<td>2.3†</td>
</tr>
<tr>
<td></td>
<td></td>
<td>β2-microglobulin 2</td>
<td>2.3†</td>
<td>3</td>
<td>1.8†</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Extrapolated sites 4 to 6</td>
<td>1.8†</td>
<td></td>
<td>0.6†</td>
</tr>
</tbody>
</table>

Abbreviations: LDH, lactate dehydrogenase; PS, performance status; GELA, Groupe d’Etude des Lymphomes de l’Adulte; TTF, time to treatment failure; IPI, International Prognostic Index; GS, overall survival.

*3-year survival.
†3-year TTF.
‡Median OS (years).

PubMed was searched for English-language articles using the following terms: “prognosis,” “outcome,” or “survival,” and “diffuse large cell lymphoma” or “intermediate-grade lymphoma.” Only articles studying biologic markers in at least 40 patients treated with anthracycline-based chemotherapy and observed for a medium of at least 12 months were selected. Studies performed on specific subtypes of extranodal lymphoma (such as CNS) and studies in human immunodeficiency virus were excluded. The bibliography of each relevant article was screened for the identification of additional references. The retrieved articles were subdivided into two categories: articles describing individual prognostic biomarkers, which are reviewed in Individual Prognostic Biomarkers; and articles proposing prognostic models based on the combination of several biomarkers,
which are summarized in sections on Gene Array and Immunohistochemistry Models. Prognostic markers that are incorporated into predictive models but for which individual predictive power was not reported separately in an article dedicated only to its prognostic value, are not discussed in Individual Prognostic Biomarkers.

**INDIVIDUAL PROGNOSTIC BIOMARKERS**

Cells in the emerging malignant clone accumulate genetic or epigenetic changes that lead to an aberrant gene activity and altered phenotypes which are subject to selection. The “hallmark features” of the cancer cell phenotype, which contribute to aggressive tumor behavior, are the capacity for sustained proliferation, evasion of apoptosis, disregard of signals to stop proliferation and to differentiate, capacity to invade and promote angiogenesis. Alteration in each of these processes might contribute to tumor resistance to therapy and serve as a useful prognostic biomarker (Fig 1). Multiple studies have tried to associate the survival of DLBCL patients with specific molecular factors underlying the “hallmark features” of the lymphoma cells.

**Cell Cycle Regulatory Molecules**

TP53. TP53 is a tumor suppressor gene that acts as a multifunctional transcription factor involved in cell cycle arrest, apoptosis, cell differentiation, replication, DNA repair, and maintenance of genomic stability. Mutations in TP53 are detected in 18% to 30% of large B-cell lymphoma patients and have been associated with shorter survival or disease-free survival (DFS) in some studies but not in others (Table 2). The variable predictive power of TP53 mutations in DLBCL might be attributed to different methodologies used for their detection. Although some studies examined the presence of TP53 mutations directly at the genomic levels; others relied on immunohistochemistry (IHC). Wild-type p53 has a short half-life and is not usually detected by immunostaining, whereas mutated p53 has prolonged half-life due to protein stabilization and thus can be detected by IHC. The correlation between TP53 mutations and immunohistochemical detection of p53, however, is not perfect.

![Schematic representation of pathophysiological events contributing to pathogenesis and outcome of diffuse large B-cell lymphoma.](image)

**Apoptotic Proteins**

Survivin. Survivin, a member of the inhibitor of apoptosis family of proteins, is expressed in a high percentage of human cancers but not in terminally differentiated adult tissues. Survivin expression has been associated with a significantly shorter 5-year OS in patients with DLBCL (Table 3). BCL2. BCL2 is an antiapoptotic protein located mainly in the inner mitochondrial membrane. It was originally discovered due to its involvement in the (t(14;18)(q32;q21)) translocation, which juxtaposes the BCL2 gene from 18q21 to immunoglobulin heavy-chain.
locus, resulting in BCL2 overexpression.\textsuperscript{50,51} However, BCL2 overexpression in DLBCL can also be caused by other mechanisms, such as amplification of the BCL2 gene.\textsuperscript{52,53} Approximately 47% to 58% of DLBCL tumors express BCL2 protein.\textsuperscript{44,46}

Because BCL2 protein is postulated to promote survival advantage on malignant B cells through inhibition of apoptosis, several studies examined the correlation among BCL2 expression, BCL2 translocations, and outcome of DLBCL patients.\textsuperscript{26,33,41,44-47,54,55} BCL2 translocation status did not correlate with survival of DLBCL. In contrast, most studies showed either inferior DFS or OS of patients whose tumors exhibit high BCL2 protein expression (Table 3).\textsuperscript{33,44-46}

Furthermore, BCL2 expression (either mRNA or protein) was also incorporated into several multigene prognostic models (see Gene Expression Profiling and DLBCL Prognostic Signatures and Immuno-histochemical Models for Prediction of DLBCL Survival). Notably, BCL2 protein expression studies used nonuniform IHC cutoffs (10% to 60%) for definition of BCL2-positive or -negative tumors (Table 3), making apparent the lack of agreement in the literature on the optimal cutoff point for BCL2 staining.

**Caspses.** A recent study evaluated whether inhibition of caspase 8 and/or caspase 9 apoptosis signaling pathways predict clinical outcome of patients with nodal DLBCL. The caspase 8 inhibition profile, assessed by immunostaining for cellular FLICE inhibitory protein (c-FLIP) and caspase 3, was associated with an excellent clinical outcome (100% CR, 9% relapse rate, 5-year OS of 100%), whereas the caspase 9 inhibition profile, assessed by expression of BCL2 and X-linked inhibitor of apoptosis, was strongly associated with poor response to chemotherapy and short survival (38% CR, 34% relapse rate, 5-year OS of 22%).\textsuperscript{56}

**B-Cell Differentiation**

The differentiation stage of B-lymphocytes at which neoplastic transformation occurs may define the biologic behavior and the outcome of DLBCL patients. DLBCL tumors demonstrate heterogeneous expression of surface and intracellular proteins and exhibit different status of ongoing mutations in immunoglobulin genes, thus suggesting different ontogenetic origin.\textsuperscript{57} Multiple studies tried to correlate individual B-cell ontogenetic markers with patient survival.

**BCL6.** The BCL6 proto-oncogene, identified by virtue of its involvement in chromosomal translocation affecting band 3q27, is the most common translocation detected in DLBCL.\textsuperscript{38-40} BCL6 functions as a sequence-specific transcriptional repressor and is necessary for

---

### Table 2. DLBCL: Cell Cycle Regulation Prognostic Biomarkers

<table>
<thead>
<tr>
<th>Author (year)</th>
<th>Marker</th>
<th>No. of Patients</th>
<th>% Positive</th>
<th>Time Period</th>
<th>Marker Positive (%)</th>
<th>Marker Negative (%)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ichikawa (1997)\textsuperscript{24}</td>
<td>p53</td>
<td>102</td>
<td>22</td>
<td>5 years</td>
<td>16</td>
<td>64</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Zhang (1999)\textsuperscript{23}</td>
<td>p53</td>
<td>158</td>
<td>58</td>
<td>5 years</td>
<td>41</td>
<td>48</td>
<td>.04</td>
</tr>
<tr>
<td>Leroy (2002)\textsuperscript{25}</td>
<td>p53</td>
<td>69</td>
<td>23</td>
<td>6 years</td>
<td>44</td>
<td>79</td>
<td>.01</td>
</tr>
<tr>
<td>Sohn (2003)\textsuperscript{26}</td>
<td>p53</td>
<td>98</td>
<td>23</td>
<td>3 years</td>
<td>43</td>
<td>65</td>
<td>NS</td>
</tr>
<tr>
<td>Maartense (2004)\textsuperscript{27}</td>
<td>p53</td>
<td>327</td>
<td>18</td>
<td>5 years</td>
<td>33</td>
<td>23</td>
<td>.47</td>
</tr>
</tbody>
</table>

**Table 3. DLBCL: Apoptosis-Related Prognostic Biomarkers**

<table>
<thead>
<tr>
<th>Author (year)</th>
<th>Marker</th>
<th>No. of Patients</th>
<th>Cutoff for Positive (%)</th>
<th>% Positive</th>
<th>Time Period</th>
<th>Marker Positive</th>
<th>Marker Negative</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adida (2000)\textsuperscript{29}</td>
<td>Survivin</td>
<td>222</td>
<td>70-90</td>
<td>60</td>
<td>5 years</td>
<td>40</td>
<td>50</td>
<td>.02</td>
</tr>
<tr>
<td>Hill (1996)\textsuperscript{24}</td>
<td>BCL2</td>
<td>135</td>
<td>10-55</td>
<td>55</td>
<td>10-year relapse-free survival</td>
<td>35</td>
<td>69</td>
<td>.03</td>
</tr>
<tr>
<td>Hermine (1996)\textsuperscript{25}</td>
<td>BCL2</td>
<td>151</td>
<td>60-45</td>
<td>45</td>
<td>3 years</td>
<td>61</td>
<td>81</td>
<td>&lt;.05</td>
</tr>
<tr>
<td>Kramer (1996)\textsuperscript{23}</td>
<td>BCL2</td>
<td>165</td>
<td>50-45</td>
<td>45</td>
<td>2 years</td>
<td>34</td>
<td>60</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Gascoyne (1997)\textsuperscript{26}</td>
<td>BCL2</td>
<td>116</td>
<td>10-52</td>
<td>24</td>
<td>3 years</td>
<td>24</td>
<td>65</td>
<td>&lt;.006</td>
</tr>
<tr>
<td>Barrans (2002)\textsuperscript{27}</td>
<td>BCL2</td>
<td>169</td>
<td>50-52</td>
<td>52</td>
<td>3 years</td>
<td>34</td>
<td>57</td>
<td>.03</td>
</tr>
<tr>
<td>Colomo (2003)\textsuperscript{21}</td>
<td>BCL2</td>
<td>126</td>
<td>NR</td>
<td>59</td>
<td>5 years</td>
<td>36</td>
<td>69</td>
<td>.06</td>
</tr>
<tr>
<td>Sohn (2003)\textsuperscript{26}</td>
<td>BCL2</td>
<td>94</td>
<td>50</td>
<td>26</td>
<td>3 years</td>
<td>41</td>
<td>66</td>
<td>.08</td>
</tr>
</tbody>
</table>

**Abbreviations:** DLBCL, diffuse large B-cell lymphoma; NR, not reported.
germinal center (GC) formation. Its expression is strictly regulated during B-cell ontogeny, so that it is almost exclusively expressed only in GC lymphocytes or lymphomas originating at the GC differentiation stage. Translocations affecting the BCL6 locus and mutations in the 5′ nontranslated regulatory region of the gene may deregulate its expression. Initial studies investigated the prognostic significance of the BCL6 rearrangements and mutations in DLBCL. Oeff et al subsequently incorporated HGAL into a multigene model for prediction of DLBCL survival (see Gene Expression Profiling and DLBCL Prognostic Signatures).

CD10. CD10, also known as the common acute lymphoblastic leukemia antigen, is a membrane-associated neutral endopeptidase. During lymphocyte differentiation, CD10 first appears on pro-B cells and is lost during maturation to naïve B cells. CD10 reappears on the cell surface during antigen-dependent GC maturation and thus serves as a marker of GC derivation. CD10 expression is detected in 20% to 30% of DLBCL tumors. Analysis of CD10 expression, a marker of GC origin, as a prognosis predictor in DLBCL yielded conflicting results. OS in patients with increased CD10 expression has been reported to be increased or not affected. Taken together with the findings that the expression of other GC-specific markers (eg, BCL6 and HGAL) correlates with DLBCL survival, these observations indicate that not every marker of GC origin has prognostic significance.

CD5. CD5 is an antigen expressed primarily by T cells and a small subset of B cells (B-1 cells). It is expressed in most patients with chronic lymphocytic leukemia and mantle-cell lymphoma, but only in 10% of patients with DLBCL. CD5-positive DLBCL tumors commonly are observed in elderly women and usually present at advanced stage, with high levels of lactate dehydrogenase that account for the high IPI score commonly reported in these patients. Although one study showed a markedly shorter survival in DLBCL patients with CD5-positive tumors, this observation was not confirmed in an additional smaller study.

FOXP1. FOXP1 is a transcription factor expressed in normal activated B cells and a subset of DLBCL with a nongerminial center phenotype. Barrans et al found a significantly inferior median OS for patients with uniform high expression of FOXP1, which was independent of IPI. This finding was confirmed by a subsequent study by Banham et al, in which FOXP1 was associated with a dramatic difference in outcome. Patients with FOXP1 expression had a significantly worse OS and a greater tendency for early progression across

### Table 4. DLBCL: B-Cell Differentiation Prognostic Biomarkers

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th>Marker</th>
<th>No. of Patients</th>
<th>% Positive</th>
<th>Time Period</th>
<th>Overall Survival</th>
<th>Marker Positive</th>
<th>Marker Negative</th>
<th>P</th>
<th>DFS, EFS, or PFS</th>
<th>Time Period</th>
<th>Marker Positive</th>
<th>Marker Negative</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lossos (2001)</td>
<td>BCL6</td>
<td>61</td>
<td>70</td>
<td>Median, months</td>
<td>171</td>
<td>24</td>
<td>.007</td>
<td></td>
<td>5-year DFS</td>
<td>55</td>
<td>15</td>
<td>.02</td>
<td></td>
</tr>
<tr>
<td>Hans (2004)</td>
<td>BCL6</td>
<td>152</td>
<td>56</td>
<td>5 years</td>
<td>69</td>
<td>30</td>
<td>&lt;.001</td>
<td></td>
<td>5-year EFS</td>
<td>57</td>
<td>35</td>
<td>.013</td>
<td></td>
</tr>
<tr>
<td>Lossos (2003)</td>
<td>HGAL</td>
<td>54</td>
<td>44</td>
<td>Median, months</td>
<td>67</td>
<td>33</td>
<td>.01</td>
<td></td>
<td>Median DFS</td>
<td>20</td>
<td>Not reached</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Onshima (2001)</td>
<td>CD10</td>
<td>138</td>
<td>29</td>
<td>5 years</td>
<td>68</td>
<td>48</td>
<td>&lt;.031</td>
<td></td>
<td>5-year DFS</td>
<td>55</td>
<td>47</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Hans (2004)</td>
<td>CD10</td>
<td>152</td>
<td>28</td>
<td>5 years</td>
<td>74</td>
<td>44</td>
<td>.019</td>
<td></td>
<td>5-year EFS</td>
<td>51</td>
<td>47</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Colombo (2003)</td>
<td>CD10</td>
<td>127</td>
<td>72</td>
<td>5 years</td>
<td>40</td>
<td>54</td>
<td>NS</td>
<td></td>
<td>5-year EFS</td>
<td>68</td>
<td>55</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Fabiani (2004)</td>
<td>CD10</td>
<td>98</td>
<td>34</td>
<td>5 years</td>
<td>68</td>
<td>65</td>
<td>NS</td>
<td></td>
<td>5-year EFS</td>
<td>68</td>
<td>55</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Yamaguchi (2002)</td>
<td>CD5</td>
<td>493</td>
<td>22</td>
<td>5 years</td>
<td>34</td>
<td>50</td>
<td>.0026</td>
<td></td>
<td>Median PFS, years</td>
<td>1.25</td>
<td>5.7</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Bahnam (2005)</td>
<td>FOXP-1</td>
<td>99</td>
<td>59</td>
<td>Median, years</td>
<td>1.6</td>
<td>12.2</td>
<td>.0001</td>
<td></td>
<td>1.25</td>
<td>5.7</td>
<td>NS</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Barrans (2004)</td>
<td>FOXP-1</td>
<td>126</td>
<td>18</td>
<td>Median, months</td>
<td>10</td>
<td>45</td>
<td>.015</td>
<td></td>
<td>1.25</td>
<td>5.7</td>
<td>NS</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Ogawa (2004)</td>
<td>CD21s</td>
<td>240</td>
<td>36</td>
<td>5 years</td>
<td>70</td>
<td>38</td>
<td>.00001</td>
<td></td>
<td>1.25</td>
<td>5.7</td>
<td>NS</td>
<td>NS</td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:** DLBCL, diffuse large B-cell lymphoma; DFS, disease-free survival; EFS, event-free survival; PFS, progression-free survival; NR, not reported; NS, not significant.
all subsets of IPI. However, the IHC cutoffs (uniform high expression and 30%, respectively) were different between these two studies. Furthermore, usage of 30% cutoff by Barrans et al on their patients led to the disappearance of survival prediction for FOXP1 protein expression.

**PKC-β.** PKC-β is a protein kinase involved in B-cell receptor signaling. It is expressed at higher levels in activated normal B lymphocytes compared with germinal center–derived cells. A recent study demonstrated that expression of this protein is a significant predictor for worse OS and DFS.

**CD21.** CD21 is an antigen expressed on both B-cells and follicular dendritic cells. CD21 expression is strongest on marginal zone B cells and activation of B cells leads to decreased CD21 expression. CD21 expression in DLBCL has been associated with improved survival in a recent study. CD21 is present in two isoforms: a short form of CD21, CD21S, which is specific for B cells and lacks exon 10a, and the long form, CD21L, which is specific for follicular dendritic cells. Evaluation of CD21S expression in DLBCL patients revealed that patients with CD21S-positive tumors (36%) had higher CR rates and a better 5-year OS rate, independent of IPI.

**Adhesion Molecules** Adhesion molecules are primarily involved in the regulation of lymphocyte homing and migration to the sites of inflammation. These molecules, in addition to their physiologic functions, may also be involved in tumor invasion and metastases. Two of the adhesion molecules have been associated with worse prognosis in patients with DLBCL, intercellular adhesion molecule-1 (ICAM-1) and CD44 (Table 5).

**ICAM-1.** ICAM-1 is a cell-surface receptor that is involved in lymphoid trafficking and extravasation. Loss of ICAM-1 expression on DLBCL cells has been associated with decreased tumor-infiltrating T lymphocytes, the number of which was reported to be predictive of relapse-free survival of DLBCL patients. Low expression of ICAM-1 in aggressive NHL was correlated with advanced stage, extranodal involvement, bone marrow infiltration, poor response to therapy, and worse survival. ICAM-1 can be shed from the cellular surface into serum, resulting in a soluble form. The presence of increased serum levels of soluble ICAM has been associated with shorter survival in both nodal and primary extranodal DLBCL.

**CD44.** CD44 is a cell surface glycoprotein that is widely distributed in different cell types and tissues. Due to alternative RNA splicing or posttranslational modifications, different isoforms of CD44 may be created. Although normal lymphocytes usually express the standard 90-kd CD44 isoform (CD44s), larger CD44 variants are preferentially expressed in epithelial cells. CD44 variants may also be found in solid tumors, activated lymphocytes, and lymphomas. Elevated levels of serum CD44 have been associated with decreased rates of complete response in both NHL and Hodgkin’s disease. Although earlier studies suggested an inferior outcome for DLBCL patients with increased levels of serum CD44, more recent studies did not confirm these observations (Table 5).

**Angiogenesis** Angiogenesis is the formation of new capillaries from existing blood vessels. This process, an essential step for tumor growth beyond a few cubic millimeters, is tightly regulated in adults by a variety of regulatory proteins that act by either stimulating or inhibiting blood vessel growth. Both proangiogenic and antiangiogenic growth factors can be found in the blood, and their serum levels may be of prognostic value.

**Endostatin.** Endostatin, a 20-kd C-terminal fragment of collagen XVIII, is a potent inhibitor of angiogenesis and tumor growth in vivo. Increased levels of serum endostatin were associated with a significantly worse outcome in a series of 60 DLBCL patients treated with anthracycline-based combination regimens.

**Vascular endothelial growth factor.** Vascular endothelial growth factor (VEGF) and basic fibroblast growth factor are mitogens for the endothelial cells and potent inducers of angiogenesis. Serum VEGF (S-VEGF) levels are increased in up to 63% of patients with DLBCL. Salven et al in their initial study did not find statistical correlation between S-VEGF levels and survival in 41 patients with DLBCL. However, in the subsequent study analyzing 78 DLBCL patients, the same authors observed longer 5-year OS in patients with S-VEGF levels below 499 pg/mL. The prognostic power of S-VEGF was improved when serum basic fibroblast growth factor was incorporated into the prognostic model. Patients with increased levels of both S-VEGF and fibroblast growth factor had a significantly shorter 5-year OS compared with patients with high levels of both growth factors (0% and 57%, respectively).

**Matrix metalloproteinase 9.** Matrix metalloproteinases (MMPs) are zinc-containing endopeptidases capable of degrading the extracellular matrix. They are broadly divided into collagenases, stromelysins, membrane-bound MMPs, and gelatinases. The latter group comprises gelatinases A (MMP-2) and B (MMP-9). Loss of the tight MMP control in malignancies...
is associated with excessive destruction of extracellular matrix, increased angiogenesis, and tumor spread.\textsuperscript{114}

Sakata et al\textsuperscript{115} recently evaluated the roles of the MMP-9 and MMP-2 as prognostic biomarkers in 75 patients with DLBCL. MMP-9 and MMP-2 were expressed in 49\% and 45\% of DLBCL tumors, respectively. In DLBCL patients with stage I and II disease, the expression of MMP-9 but not of MMP-2 was associated significantly with inferior 5-year survival (58\% vs 83\%; \( P < .01 \)). However, a significant proportion of the analyzed patients were treated by radiotherapy alone without the standard anthracycline-containing chemotherapy, thus markedly confounding the interpretation of the reported data.\textsuperscript{115}

**Other Single-Molecule Prognostic Biomarkers**

\textit{nm23-H1}. The \textit{nm23-H1} protein, a differentiation inhibitory factor involved in tumor metastasis, has been identified as a prognostic biomarker in DLBCL. Both increased serum levels\textsuperscript{116} and intensity factor involved in tumor metastasis, has been identified as a prognostic biomarker in DLBCL. Both increased serum levels\textsuperscript{116} and intensity

\textit{Hepatocyte growth factor}. Hepatocyte growth factor and its receptor \textit{c-MET} are coexpressed in several epithelial tumors. Patients with DLBCL and low-risk IPI whose tumors express hepatocyte growth factor or \textit{c-MET} have been found to have worse prognosis. Neither marker was predictive for outcome in patients with high-risk IPI.\textsuperscript{122}

\textit{Major histocompatibility complex molecules}. Infiltrating T lymphocytes isolated from B-cell lymphomas may recognize specific epitopes of the malignant clone.\textsuperscript{123} It was reported that the number of infiltrating T cells in the initial lymphoma biopsy from DLBCL patients was predictive of relapse-free survival.\textsuperscript{96} Loss of major histocompatibility complex (MHC) class I and class II (HLA-DP and HLA-DR) expression was reported to correlate with shortened relapse-free survival and OS.\textsuperscript{124,125} Fewer CD\textsuperscript{8} T cells were detected in MHC class II–negative patients compared with MHC class II–positive patients; thus supporting the hypothesis that loss of tumor immunosurveillance might contribute to the inferior outcome of DLBCL patients.\textsuperscript{125} In contrast, the presence of a high percentage of activated cytotoxic T-cell lymphocytes (CTLs; defined as CD3\textsuperscript{+} tumor-infiltrating lymphocytes with granzyme B expression) in biopsy specimens from patients with primary nodal DLBCL has been associated with worse outcome, independent of IPI. In this study, MHC-I–positive tumors had a higher percentage of CTLs, suggesting that the number of activated CTLs depends on the intact MHC-I–dependent antigen presentation. Nevertheless, despite the correlation with CTLs, MHC-I expression did not have prognostic value.\textsuperscript{126}

Individual biomarker studies provided valuable additional prognostic information to the IPI and have increased our current knowledge on the pathophysiologic mechanisms involved in lymphomagenesis. However, a large number of studies have yielded conflicting and nonconclusive results. Both might be caused by inherent problems associated with retrospective nature of most of the published studies. In addition, the limited role in clinical practice of individual biomarkers as reliable predictors of prognosis might be a consequence of the complexity of biologic processes, with the involvement of multiple genes, signaling pathways, and regulatory mechanisms, making single genes or molecules unable to reflect the heterogeneity of malignant cells accurately.

Therefore, more recent studies have explored the relationship between the DLBCL prognosis and the molecular features of the tumors using genome-scale expression profiles assessed by DNA microarrays.\textsuperscript{79,87,127} or real-time polymerase chain reaction (PCR).\textsuperscript{38} DNA microarrays are created in two basic forms, including DNA deposition (spotted arrays) on glass slides or by in situ synthesis of oligonucleotide arrays. Affymetrix arrays (Affymetrix Inc, Santa Clara, CA) are generated by in situ synthesis of short oligonucleotide fragments (20 to 24 nucleotides) of single-stranded DNA by photolithography. Spotted DNA microarrays are made by mechanical deposition of PCR-amplified double-stranded cDNA fragments (approximately 500 base pairs) or long oligonucleotides (50 to 70 mer) arrayed as small spots on glass slides. Gene expression profiling studies of DLBCL used Affymetrix arrays, spotted cDNA arrays (Lymphochip), or quantitative real-time PCR. The Lymphochip was designed with the selection of genes expressed preferentially in lymphoid cells and genes with known or suspected roles in cancer biology.\textsuperscript{128} A comparison of the different platforms is listed in Table 6.

Specific alterations occurring in the tumor cells modify their pattern of mRNA expression, characterizing their molecular signature or fingerprint, and influence their characteristics and behavior. The analytic approaches used for analysis of array data can be subdivided into two categories: unsupervised and supervised. The unsupervised approach analyzes the array data without using external information such as clinical parameters or survival time. In contrast, the supervised approach aims to identify genes, the expression of which correlates with some external variables. When both the unsupervised and supervised methods were used, array studies of DLBCL found gene expression signatures that correlated significantly with clinical outcome.\textsuperscript{129}

The pivotal study of microarrays in DLBCL was performed with the use of a cDNA Lymphochip array.\textsuperscript{85} Using an unsupervised hierarchichal clustering learning method in tumor samples from 42 DLBCL patients treated with anthracycline-based chemotherapy, Alizadeh et al\textsuperscript{87} identified two distinct DLBCL subgroups based on the expression
pattern of genes characteristic of GC B cells (GC signature genes) and in vitro activated peripheral-blood cells (activated B-cell signature genes). DLBCL subtypes expressing these two gene expression signatures were named GCB-like DLBCL and activated B-cell (ABC)-like DLBCL, respectively. These two DLBCL subgroups had a significantly different outcome, with 76% of GCB-like patients alive at 5 years compared with 16% of ABC-like DLBCL patients (P < .01). This improved survival for patients with GCB-like DLBCL remained statistically significant even when only patients with low-risk disease (IPI 0 to 2) were evaluated. Notably, some of the genes comprising the GC B-like and ABC-like signatures were among the single prognostic markers in DLBCL, such as BCL6 and HGAL in the former and BCL2 in the latter. These observations were further confirmed in a larger study performed by the Lymphoma and Leukemia Molecular Profiling Project (LLMPP) group, which analyzed gene expression profiles in 240 DLBCL patients treated with CHOP-like regimens. In this study, the 5-year OS of patients with GCB-like and ABC-like DLBCL was 60% and 35%, respectively. 79

Although expression profiling by DNA microarrays established the presence of biologically distinct subtypes of DLBCL, the unsupervised analytic methods used in these studies were unable to indicate the relative contribution of individual genes or to construct a survival prediction model based on a relatively small number of genes that might easily be applied in routine practice. Supervised analysis of gene expression data was used to address these questions. Shipp et al 127 derived a 13-gene (dystrophin related protein 2, protein kinase C gamma, MINOR, 5-hydroxytryptamine 2B receptor, H731, transducin-like enhancer protein 1, PDE4B, protein kinase C-beta-1, oviductal glycoprotein, zinc-finger protein C2H2-150, and three expressed sequence tags), IPI-independent predictive model from a cohort of 58 patients whose lymphomas were analyzed by Affymetrix oligonucleotide microarrays. Only three of these 13 genes were present in the data set reported by Alizadeh et al, 87 and on re-evaluation of their predictive power in this data set, only two were associated with survival (Lossos, unpublished observation). Rosenwald et al 79 also used supervised analysis of 160 DLBCL patients to derive a predictive model based on the expression of 17 genes and then validated this model in a separate data set of 80 DLBCL. The genes that comprised this model were derived from four gene expression signatures (GC B cell, MHC class II, lymph node, and proliferation) and included BCL6, HGAL, clone 1334260, HLA-Dpα, HLA-Dqα, HLA-Drα, HLA-DRβ, α-actinin, collagen type III α1, connective-tissue growth factor, fibronectin 1, KIAA0233, urokinase plasminogen activator, c-myc, E21G3, NPM3, and BMP6. Of interest, there was no overlap of genes in the Shipp 127 and the Rosenwald 79 models. Although the disparity between the studies was initially attributed to technical differences and to the composition of the arrays, recent studies on primary mediastinal lymphoma using the same two platforms, (Lymphochip and Affymetrix oligonucleotide) resulted in similar findings. 130 It remains possible, however, that different criteria for patient selection and statistical algorithms used for the construction of the predictive models could be responsible for this disparity. Analysis of the LLMPP group data by a different statistical method for identifying prognostic genes (significance analysis of microarrays), 131 suggested that some but not all of the genes incorporated into this survival model had predictive power. 38 Certainly every predictive model needs to be validated on an independent cohort of patients to confirm that it works generally and not just on the group of patients from which it was derived. 132,133

To resolve this disparity, Wright et al 80 recently designed a method based on Bayes’ rule to classify DLBCL tumors as GCB-like DLBCL characterized by prolonged survival and ABC-like DLBCL characterized by shorter survival. For this classification the authors used expression data from 14 genes identified within the LLMPP group data set that were also analyzed by Affymetrix arrays by Shipp et al. They used similar criteria for patient selection and statistical algorithms used for the construction of the predictive models. However, to apply this method to the Affymetrix data, shifting and scaling manipulations of expression values were required to achieve matching mean and variance levels, a manipulation that may limit the widespread application of this method in clinical practice.

In an attempt to devise a technically simple model that would be applicable for routine clinical use, we recently evaluated the expression of mRNA for 36 candidate genes reported to predict survival in DLBCL. The expression of these 36 genes was measured in DLBCL tumors of 66 patients treated with anthracycline-based regimens. The top six genes ranked on the basis of their predictive power in univariate analysis were used to construct a model based on the relative contribution of each of them in a multivariate analysis, where the LMO2, BCL6, and FNI genes predicted a longer survival and the CCND2, SCTXA3, and BCL2 genes predicted a short survival. On the basis of this

| Table 6. Comparison of Gene Expression Platforms Used in DLBCL Studies |
|-----------------|-----------------|-----------------|
| Characteristic   | Affymetrix      | Lymphochip      | Real-Time TaqMan PCR |
| Platform         | Commercially available | Custom spotted PCR amplified cDNA array | Commercially available primer/probe preloaded low density arrays |
| No. of genes tested | Thousands | Thousands | Tens |
| Method           | Single-color hybridization | Double-color hybridization | Real-time PCR |
| Target specificity | High | Good* | Very high |
| Relative dynamic range of expression | Intermediate | Lowest | Highest |
| RNA source       | Frozen/fresh tissue | Frozen/fresh tissue | Frozen/fresh tissue (potentially paraffin-derived tissue) |
| RNA quantity     | 5-10 μg tRNA | 5-10 μg tRNA | 20-30 ng/gene |

Abbreviations: DLBCL, diffuse large B-cell lymphoma; PCR, polymerase chain reaction; *
Not depends on the number of clones with verified sequence.
model, patients could be subdivided into low-, intermediate-, and high-risk groups. Both the 5-year survival (65% in the low-risk group, 49% in the intermediate-risk group, and 15% in the high-risk group) and the mean survival time (8.7, 7.1, and 3.8 years, respectively) were significantly different among the DLBCL subgroups identified by this model. This model was independent of the IPI and added to its predictive power. In addition, it was also able to predict survival when applied to published microarray gene-expression data from previous studies that used either Lymphochip or Affymetrix oligonucleotides arrays. The six-gene model could identify approximately one third of all patients with DLBCL with 5-year survival of less than 27% who may require a different therapeutic approach and may benefit from investigational therapeutic modalities.

Examination of mRNA expression by array or real-time reverse transcriptase PCR technology provides quantitative measurement of gene expression but requires fresh vital tissues or fresh frozen specimens, thus limiting applicability of this method in routine practice in the community. In contrast, IHC of paraffin-embedded tissue provides only semiquantitative assessment of protein expression but is used routinely for histologic evaluation of lymphoma specimens. Tissue microarrays (TMAs), which allow high-throughput protein expression studies, may facilitate the identification of prognostic IHC models for immediate clinical application.

Several recent studies have focused on the analysis of protein expression of selected markers by IHC studies in patients with DLBCL in an effort to define immunophenotypic profiles that better identify risk groups. Protein expression studies in DLBCL, however, have yielded conflicting results. Hans et al demonstrated that IHC with a combination of CD10, BCL6, and MUM1 antibodies could classify DLBCL patients as long- and short-term survivors, and had a positive predictive value of 87% and 73% for correct classification of lymphomas as GCB-like and ABC-like DLBCL, respectively, when considering cDNA microarray classification as the gold standard. Furthermore, this study demonstrated that individual expression of CD10 and BCL6 was associated with prolonged OS, whereas individual expression of MUM1, BCL2, and cyclin D2 was associated with shorter OS, thus confirming some of the previous studies detailed in Tables 3 and 4. Barrans et al reported that IHC with antibodies to GC markers (CD10 and BCL6) and to BCL2 in combination with the IPI could improve risk stratification in DLBCL. However, these results are not concordant with the data from Colomo et al. Saez et al analyzed the expression of 52 proteins in DLBCL and constructed a different model derived from logistic regression analysis and based on expression of eight markers (cyclin E, CDK1, CDK2, SKP2, EBER, MUM1, Rb-P, and BCL6).

Current data demonstrate that there is no consensus regarding the best IHC model to describe the heterogeneity of DLBCL. Nevertheless, identified candidate proteins with known contribution to survival prediction should be examined further in future attempts to construct IHC-based survival models.

The purpose of either individual biomarkers or pattern-based biomarkers models is to provide a basis for predicting survival, choice of initial treatment, stratification of patients in clinical trials, accurate communication among healthcare providers, and uniform reporting of outcomes. Useful biomarkers for prediction of survival of DLBCL patients must demonstrate clinical value complementary to the IPI. A number of such individual biomarkers that contribute to the survival prediction by IPI have been identified. Furthermore, the use of DNA microarray methodology allowed the development of new classifications based on molecular profiling, with the description of specific disease entities sharing similar biologic features, clinical behavior, and outcome. A molecular classification of DLBCL currently is a reality that is here to stay. It allowed construction of RNA and protein survival prediction models based on expression pattern of a relatively small number of genes or proteins. However, the prime time for their incorporation into routine clinical practice has not yet arrived. The reasons for this are as follows.

First, validation is lacking. For the incorporation of individual prognostic biomarkers or biomarker models into routine clinical practice, each scientifically vetted test must pass several well-controlled evaluation steps aimed to assess the robustness and reproducibility of the assay, examine the predictive potential of the biomarker in retrospective longitudinal cohorts of patients, and confirm its validity in prospective well-designed studies. None of the proposed prognostic biomarkers has successfully passed all these obstacles. Most of the reported DLBCL prognostic biomarkers and models were derived from retrospective studies that used available tissue specimens and serum samples that were not collected and handled uniformly. The absence of uniform frozen specimen collections and handling as well as common nonuniform RNA extraction protocols might significantly compromise the conclusions of the RNA-based models. Moreover, almost all retrospective studies were carried out in the absence of a predetermined written protocol, eligibility criteria, primary end point, or predetermined statistical analysis, and commonly included too few patients. These pitfalls might contribute to the contradictory results of some of the proposed prognostic markers or models as described, and should be accounted for in the design of prospective studies for biomarkers validation.

Although internal validation on subsets of original patient cohort is commonly performed, frequently as an essential part of prognostic model development and refinement, this process does not eliminate the inherent biases associated with uncontrolled selection of retrospectively available specimens that may lead to model over fitness in the tested study group, and thus cannot and should not replace the need for formal external validation in independent groups of patients. Such validation in independent, well-designed, large prospective studies is of particular importance, given that it will confirm that the suggested methodologies to assess the expression of prognostic markers are robust, reproducible, and are independent of potential variability in sample handling or applied methodology. In addition, these prospective studies will also serve for external validation of method reproducibility in different laboratories.

For confirmation of the proposed RNA-based prognostic models in a prospective setting, it is imperative to attribute specific attention to uniformity and unbiased collection and handling of the representative
BCL2 expression was more pronounced in the patients treated with BCL2. The results indicate that the negative prognostic impact of phomes de l'Adulte recently revisited the prognostic significance of and CHOP (R-CHOP). In view of the improved survival achieved at the treatment, whereas the standard treatment quantitative and observer-independent model.

In addition, development of robust and reproducible prediction models based on RNA extracted from paraffin-embedded tissues using, for example, real-time PCR, would be of potential value. Similar to IHC methodology, it would use widely available paraffin-embedded tissues and would eliminate the need for collection of fresh or frozen specimens. However, in contrast to IHC, it would be a more quantitative and observer-independent model.

Second, the uniqueness of prognostic biomarkers is dependent on the effectiveness of the treatment, whereas the standard treatment of DLBCL has evolved to include the anti-B-cell antibody rituximab and CHOP (R-CHOP). In view of the improved survival achieved at present by the addition of rituximab, the Groupe d’Etude des Lymphomes de l’Adulte recently revisited the prognostic significance of BCL2. The results indicate that the negative prognostic impact of BCL2 expression was more pronounced in the patients treated with CHOP therapy than in those treated with R-CHOP.134 However, this finding regarding the BCL2 expression was not confirmed in an analysis of CHOP and R-CHOP patients in the US Intergroup study.40 Although presently available data do not demonstrate conclusively that the addition of rituximab alters previously established molecular prognostic markers, it is imperative to re-examine prognostic models prospectively when therapeutic advances improve clinical outcomes. Clinical trials incorporating gene arrays and real-time reverse transcriptase–PCR are underway (B. Cheson [Cancer and Leukemia Group B] and I. Lossos, personal communication, November 2005) and promise to validate prospectively and compare the prognostic models in patients in whom rituximab is an integral part of the therapeutic regimen.

Third, to accomplish these goals and to be in a position in which existing or new prognostic models can be tested and validated easily, there is strong need to collect frozen and paraffin-embedded material that can be used for RNA extraction and construction of TMAs, respectively. Such materials should be gathered as an integral part of any planned therapeutic clinical trial.

It is clear that new high-throughput genomic technologies have yielded many potential biomarkers and biomarker models for prediction of survival in DLBCL patients. The potential of these prognostic factors is enormous and it is possible that biomarker models, used in combination with clinical factors, will become an integral part of our daily practice. This will open the way to a more individualized or patient-tailored practice of oncology and possibly will lead to identification of new specific therapeutic targets. However, multiple hurdles, discussed herein, need to be passed before prognostic biomarkers in DLBCL will become a reality.

REFERENCES

Prognostic Markers in DLBCL


---

**Authors’ Disclosures of Potential Conflicts of Interest**

The authors indicated no potential conflicts of interest.

**Author Contributions**

- **Conception and design:** Izidore S. Lossos
- **Collection and assembly of data:** Izidore S. Lossos, Daniel Morgensztern
- **Data analysis and interpretation:** Izidore S. Lossos, Daniel Morgensztern
- **Manuscript writing:** Izidore S. Lossos, Daniel Morgensztern
- **Final approval of manuscript:** Izidore S. Lossos, Daniel Morgensztern