Prognostic Value of Cellular Proliferation and Histologic Grade in Follicular Lymphoma

By Anita R. Martin, Dennis D. Weisenburger, Wing C. Chan, Elizabeth I. Ruby, James R. Anderson, Julie M. Vose, Philip J. Bierman, Martin A. Bast, Donald T. Daley, and James O. Armitage

The clinical usefulness of histologic grading in follicular lymphoma (FL) is controversial and is further compromised by the subjective nature and poor reproducibility of most systems in current use. Therefore, we decided to objectively evaluate the importance of cellular proliferation in FL, along with the current grading systems. We studied 106 patients with FL who were uniformly staged and aggressively treated. A proliferative index (PI) was determined quantitatively using an automated image analyzer and a new Ki-67 antibody that stains archival paraffin tissues. The cases were also subclassified according to the Berard, Rappaport, Luke-Collins, and Jaffe methods, and survival analysis was performed. Patients with a low PI (<40%) had a significantly longer overall survival (OS) than those with a high PI (>40%), but the PI did not predict failure-free survival (FFS). The mean PI correlated well with the subgroups in each of the various classifications. All four of the classification methods were predictive of OS, but only the Berard method appeared to predict FFS and suggest that a proportion of patients with FL may be curable. In multivariate analysis, histologic classification was the only independent predictor of OS (Berard method: relative risk, 3.1) and the International Prognostic Index was the only independent predictor of FFS (relative risk, 2.3). We conclude that the Berard method for grading of FL is clinically useful and, along with the International Prognostic Index, should be included in future clinical studies of FL. The measurement of cellular proliferation does not appear to add additional useful information in FL.

© 1995 by The American Society of Hematology.

From the Departments of Pathology and Microbiology, Preventive and Societal Medicine, and Internal Medicine, University of Nebraska Medical Center, Omaha, NE.

Submitted September 12, 1994; accepted January 31, 1995.

Supported in part by US Public Health Service Grant No. CA36727 awarded by the National Cancer Institute, Department of Health and Human Services.

Address reprint requests to Dennis D. Weisenburger, MD, Department of Pathology and Microbiology, University of Nebraska Medical Center, 600 S 42nd St, Omaha, NE 68198-3135.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. section 1734 solely to indicate this fact.

© 1995 by The American Society of Hematology.

0006-4971/95/8512-0053$3.00/0

Blood, Vol 85, No 12 (June 15), 1996; pp 3671-3679
Jaffe modification of the Berard method
Berard Method (used by the NLSG)
- FSC: ≤5 large noncleaved cells/follicle
- FM: ≥5 but <15 large noncleaved cells/follicle
- FLC: ≥15 large noncleaved cells/follicle

Rappaport method
- FSC (PDL): <25% large cells/follicle
- FM (Mixed): 24–49% large cells/follicle
- FLC (Histiocytic): ≥50% large cells/follicle

Lukes-Collins method
- FSC: <25% large noncleaved cells/follicle
- FLC: ≥25% large noncleaved cells/follicle

Jaffe modification of the Berard method
- FSC: <5 large noncleaved cells/follicle
- FM: ≥5 large noncleaved cells/follicle
- FLC: ≥50% large cells/follicle (predominance of large noncleaved and/or small noncleaved cells)

Abbreviations: hpf, high-power field (0.158 mm²); PDL, poorly differentiated lymphocytic.

Lukes-Collins methods and using a modified version of the Berard method recently described by Jaffe et al. 

Five-micron thick paraffin sections of formalin-fixed tissue were deparaffinized and rehydrated. The slides were incubated with polyclonal rabbit antihuman Ki-67 antibody (A 047 diluted 1:50; Dako, Carpinteria, CA), stained with the avidin-biotin peroxidase complex technique, and developed with H₂O₂ and diaminobenzidine (Sigma Chemical, St Louis, MO), and the nuclei were counterstained with 0.2% ethyl green in sodium acetate buffer, pH 4.0.

Quantitative image analysis. Quantitation of Ki-67 antibody staining was performed with the Cell Analysis Systems (CAS) 200 image analyzer using the CAS Quantitative Proliferation Software (Becton Dickinson, Elmhurst, IL) to obtain the percent nuclear area positive (PNA) and the cellular proliferation index (PI). These methods have been previously detailed and validated. Fifteen neoplastic follicular fields were analyzed in each case and the fields were selected from follicles with the greatest number of stained cells. Positive staining was identified as granular nuclear stain (Fig 1).

The slides were incubated with polyclonal rabbit antihuman Ki-67 antibody (A 047 diluted 1:50; Dako, Carpinteria, CA), stained with the avidin-biotin peroxidase complex technique, and developed with 

**Table 1. Histologic Methods for Grading FL**

<table>
<thead>
<tr>
<th>Method</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berard Method</td>
<td>FSC ≤5 large noncleaved cells/follicle</td>
</tr>
<tr>
<td></td>
<td>FM ≥5 but &lt;15 large noncleaved cells/follicle</td>
</tr>
<tr>
<td></td>
<td>FLC ≥15 large noncleaved cells/follicle</td>
</tr>
<tr>
<td>Rappaport method</td>
<td>FSC (PDL) &lt;25% large cells/follicle</td>
</tr>
<tr>
<td></td>
<td>FM (Mixed) 24–49% large cells/follicle</td>
</tr>
<tr>
<td></td>
<td>FLC (Histiocytic) ≥50% large cells/follicle</td>
</tr>
<tr>
<td>Lukes-Collins method</td>
<td>FSC &lt;25% large noncleaved cells/follicle</td>
</tr>
<tr>
<td></td>
<td>FLC ≥25% large noncleaved cells/follicle</td>
</tr>
<tr>
<td>Jaffe modification of Berard method</td>
<td>FSC &lt;5 large noncleaved cells/follicle</td>
</tr>
<tr>
<td></td>
<td>FM ≥5 large noncleaved cells/follicle</td>
</tr>
<tr>
<td></td>
<td>FLC ≥50% large cells/follicle (predominance of large noncleaved and/or small noncleaved cells)</td>
</tr>
</tbody>
</table>

Abbreviations: hpf, high-power field (0.158 mm²); PDL, poorly differentiated lymphocytic.

Immunostaining with Ki-67 antibody. Five-micron thick paraffin sections of formalin-fixed tissue were deparaffinized and rehydrated. To enhance antigen retrieval, sections were subjected to microwave incubation in citrate buffer at pH 6.0 twice for 5 minutes at full power and were then allowed to incubate in hot buffer for 10 minutes.

The slides were incubated with polyclonal rabbit antihuman Ki-67 antibody (A 047 diluted 1:50; Dako, Carpinteria, CA), stained with the avidin-biotin peroxidase complex technique, and developed with 

**Fig 1.** (A) Follicular lymphoma showing prominent delineation of the neoplastic follicles with the Ki-67 antibody stain. (B) Positive nuclear staining marks the cells in the G₁, S, G₂, and M phases of the cell cycle (immunoperoxidase stains; A, ×25; B, ×150.)
HISTOLOGIC GRADE IN FOLLICULAR LYMPHOMA

RESULTS

Clinical features. The clinical characteristics of the entire group are summarized in Table 2. The median age of the patients was 61 years (range, 23 to 84 years) and the median follow-up period was 46 months (range, 2 to 136 months). The predicted 5-year OS was 67% and the 5-year FFS was 39% (Fig 2). By univariate analysis, the only clinical features that were predictive for FFS included the stage of disease (P = .022), symptom status (P = .014), number of involved extranodal sites (P = .011), Karnofsky score (P = .034), and the International Prognostic Index (P = .074). The clinical features that were predictive for FFS included the stage of disease (P = .011), symptom status (P = .034), number of involved extranodal sites (P = .011), Karnofsky score (P = .022), and the International Prognostic Index (P = .014).

Pathologic features. The effect of the PI on OS is shown in Fig 3. Patients with a low PI (<40%) had a significantly longer OS than those with a high PI (≥40%; P = .022). However, the PI did not predict for FFS, and the PNA was not predictive of either OS or FFS. The number of cases by histologic subgroup for each of the classifications is shown in Table 3. When the FSC and FM categories were combined, the cases in this subgroup in our series were identical when using either the Rappaport or Jaffe methods. Each of the methods categorized the cases quite differently, with the Berard method having the most FLC cases and the Rappaport or Jaffe methods having the fewest FLC cases. The mean PI correlated well with the subgroups in each of the classifications (Table 3).

All of the classification methods were predictive of OS (Fig 4), but only the Berard method appeared to predict FFS (Fig 5). Interestingly, using the Berard method, patients with the FSC/FM type had a better OS, but poorer FFS, than those with the FLC type (Figs 4A and 5A). The apparent plateau in the FLC FFS curve (Fig 5A) suggests that a proportion of patients with this type may be cured. The predicted 8-year FFS of patients with FLC was 47% (95% confidence interval [CI], 33% to 61%) versus only 17% for those with FSC/FM (95% CI, 2% to 32%).

The superiority of the Berard method as compared with the other methods for predicting FFS is shown in Fig 6. Cases classified as FM by the Jaffe method but as FLC by the Berard method, ie, those cases with 15 or more large cells but less than 50% large cells per follicle, had the same FFS as those classified as FLC by both methods (ie, cases with ≥50% large cells per follicle). In contrast, the FFS of cases classified as FM by both methods, ie, cases with 5 to 14 large cells per follicle, fails to plateau and is characterized by a pattern of continuous relapse.

Multivariate analysis. A Cox proportional hazards model for survival analysis was used to determine which prognostic factors were predictive of OS and FFS when controlling for other factors found to be significant in the univariate analysis. For OS, only the histologic classification was an independent predictor. The Berard method was the best predictor of OS with a relative risk of death of 3.1 for those with FLC as compared with those with FSC/FM, whereas the Rappaport, Jaffe, and the Luke-Collins methods gave relative risks of 2.4, 2.4, and 2.1, respectively. Only the International Prognostic Index was an independent predictor of FFS, with a relative risk of failure of 2.3 for the high risk group (Fig 7).

DISCUSSION

The subjective nature and poor reproducibility of most histologic grading systems of FL,110 as well as ongoing controversy regarding the clinical value of such systems,1-8 led us to study our cohort of patients with FL who were uniformly staged and aggressively treated. Because a number of studies have indicated that proliferative activity is an important prognostic indicator in NHL,12-17,21-24,29 we decided to evaluate this parameter in an objective and reproducible manner. Image analysis is an ideal method for assessing proliferation in FL because it is rapid, objective, and reproducible,25,27,35,36 and the analysis can be limited to the neoplastic follicles with the greatest number of large cells. In our study, we used the automated, quantitative CAS image analysis system25,27,34-36 and a new Ki-67 antibody that works in paraffin tissue,33 thus allowing the study of archival tis-
sues. Previous studies of NHL have used a Ki-67 antibody that only works in snap-frozen tissue, thus limiting the number of cases available for study. However, one obstacle to the use of archival tissues is the variation in Ki-67 antibody staining of tissues that are fixed and processed at different institutions, and a number of cases had to be excluded from our study because of inadequate staining. The results that we obtained for FL with the CAS system are similar to those reported by Schwartz et al, who also used the same system and methods. However, our results are difficult to compare with other reported results obtained by manual methods because only a small number of cases of FL have been so studied and the Ki-67 immunohistochemical, cell counting, and histologic classification methods used were quite heterogeneous. In four such studies, the manual results were similar to our results, whereas in three other studies, two of which also included diffuse counterparts of FL, the results were somewhat lower. However, our study is the first to evaluate Ki-67 proliferative activity in FL and relate this parameter to survival.

Similar to the study of Macartney et al, which evaluated the S-phase fraction of low-grade FL by DNA flow cytometry, we found that patients with FL having a low PI had a significantly longer OS than those with a high PI (Fig 3). However, in our study, the PI was not predictive of FFS and was not an independent predictor of OS in the multivariate analysis. As in other smaller studies, we found that proliferation correlated with the histologic subgroups of FL in each of the various classifications (Table 3), thus confirming the association of cellular proliferation with large-cell cytology in NHL. In contrast, Weiss et al failed to find such a correlation in FL when using a manual, semiquantitative method to estimate Ki-67 proliferation rates. Interestingly, two studies have failed to show a correlation between the mitotic rate and survival in FL whereas a recent study found that abundant mitoses predicted for better survival in FLC NHL. However, our study indicates that the Ki-67 antibody can be used to provide important information with regard to the OS of patients with FL.

Because of the subjective nature and poor reproducibility of methods for grading FL that require estimation of the percentage of large cells, we adopted the cell counting method of Mann and Berard some years ago for use by the NLSG. Nathwani et al found that it was difficult for hematopathologists to reproducibly separate favorable from unfavorable FLs using subjective morphologic criteria alone and suggested that the cell counting method is superior to the estimation of percentages of large cells for grading FL. Although reproducibility may also be a problem with the Berard method, we have found this method to be less subjec-

<table>
<thead>
<tr>
<th>Classification</th>
<th>No. of Patients</th>
<th>Mean PI + SE (%)</th>
<th>P*</th>
<th>PI Range (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berard (NLSG) method</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FSC + FM</td>
<td>42</td>
<td>32.1 ± 1.9</td>
<td>&lt;.0001</td>
<td>8.6-59.5</td>
</tr>
<tr>
<td>FLC</td>
<td>64</td>
<td>43.4 ± 1.5</td>
<td>.0002</td>
<td>10.2-62.3</td>
</tr>
<tr>
<td>Rappaport/Jaffe methods</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FSC + FM</td>
<td>85</td>
<td>26.6 ± 1.4</td>
<td>.0002</td>
<td>8.6-59.5</td>
</tr>
<tr>
<td>FLC</td>
<td>21</td>
<td>48.3 ± 2.0</td>
<td></td>
<td>34.2-62.3</td>
</tr>
<tr>
<td>Lukes-Collins method</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FSC</td>
<td>64</td>
<td>35.3 ± 1.7</td>
<td>.0005</td>
<td>8.6-59.0</td>
</tr>
<tr>
<td>FLC</td>
<td>42</td>
<td>44.3 ± 1.6</td>
<td></td>
<td>23.2-62.3</td>
</tr>
</tbody>
</table>

Abbreviation: SE, standard error.
* P value based on the Student’s t-test.
tive and more reproducible than the other methods in our daily practice. However, when the Berard method is used, the proportion of cases in the various FL subgroups is different from that found by estimating percentages of large cells (Table 3). In our study, which was biased toward the inclusion of FLC cases (see Materials and Methods), the majority of FLs were classified as FLC type (60%) using the Berard method, whereas lesser numbers of cases were classified as such using the Lukes-Collins or Rappaport/Jaffe methods (40% and 20%, respectively). In a large and unselected survey of cases in the population-based registry of the NLSG, we found that 16% of our cases of FL were classified as

![Graph A](image1)

**Fig 4. OS of patients with FL according to the various classification systems.**

![Graph B](image2)

![Graph C](image3)

**Fig 5. FFS of patients with FL according to the various classification systems.**

FSC, 37% as FM, and 47% as FLC type using the Berard method. Our more recent data is as follows: FSC 25%, FM 36%, and FLC 39%. These findings are roughly the reverse of what has been reported using more traditional methods. In the only study similar to ours, Nathwani et al categorized FL according to the Berard method and found the following: FSC 36%, FM 42%, and FLC 22%. However, these findings are not comparable to ours. The possibility that an excess of FLC NHL occurs in Nebraska, possibly because of heavy pesticide use, is also a consideration and will be the subject of a future study. However,
these differing findings may have particular relevance to the comparison and interpretation of clinical studies that use different criteria for the selection and grading of FL.

In our study, we found that all of the classification methods were useful for predicting OS in FL (Fig 4). In fact, histologic classification was the only independent predictor of OS in the multivariate analysis, with the Berard method being the best predictor of OS (relative risk, 3.1). The Berard method also appeared to predict for FFS, whereas the other grading methods were not predictive of FFS (Fig 5). The apparent plateau in the FLC FFS curve (Fig 5A) suggests that a proportion of patients with this type of FL may be curable, as has been previously suggested by others using the Rappaport method, although longer follow-up of our series will be necessary to confirm this finding. However, using the Berard method, Bartlett et al have recently reported results similar to ours with regard to prolonged FFS in aggressively treated patients with FLC NHL and have suggested that such patients should be treated with curative intent.

The superiority of the Berard method over the other methods for predicting FFS is further evident in Fig 6. It is apparent from these curves that some cases classified as FM type using the Jaffe method and other methods, ie, those cases with 15 or more large cells but less than 50% large cells per follicle, have FFS that is identical to those with 50% or more large cells per follicle. In other words, such cases should probably be placed for clinical purposes in the FLC category, as with the Berard method. This same finding was also recently reported by Bartlett et al and indicates that some cases classified as FM type using the other methods may be curable. Further studies with longer follow-up will also be necessary to confirm this finding. However, the results of our study suggest to us that the Berard method yields more clinically relevant information than the other grading methods for FL and that it should be used along with other methods in future clinical studies.

In our study, we found that the International Prognostic Index, which was developed as a predictive model for different criteria for the selection and grading of FL.

In our study, we found that all of the classification methods were useful for predicting OS in FL (Fig 4). In fact, histologic classification was the only independent predictor of OS in the multivariate analysis, with the Berard method being the best predictor of OS (relative risk, 3.1). The Berard method also appeared to predict for FFS, whereas the other grading methods were not predictive of FFS (Fig 5). The apparent plateau in the FLC FFS curve (Fig 5A) suggests that a proportion of patients with this type of FL may be curable, as has been previously suggested by others using the Rappaport method, although longer follow-up of our series will be necessary to confirm this finding. However, using the Berard method, Bartlett et al have recently reported results similar to ours with regard to prolonged FFS in aggressively treated patients with FLC NHL and have suggested that such patients should be treated with curative intent.

The superiority of the Berard method over the other methods for predicting FFS is further evident in Fig 6. It is apparent from these curves that some cases classified as FM type using the Jaffe method and other methods, ie, those cases with 15 or more large cells but less than 50% large cells per follicle, have FFS that is identical to those with 50% or more large cells per follicle. In other words, such cases should probably be placed for clinical purposes in the FLC category, as with the Berard method. This same finding was also recently reported by Bartlett et al and indicates that some cases classified as FM type using the other methods may be curable. Further studies with longer follow-up will also be necessary to confirm this finding. However, the results of our study suggest to us that the Berard method yields more clinically relevant information than the other grading methods for FL and that it should be used along with other methods in future clinical studies.

In our study, we found that the International Prognostic Index, which was developed as a predictive model for diffuse aggressive NHL, predicts for OS and FFS in FL. Others have also recently reported that the International Prognostic Index is predictive of survival in FL, with one exception. In our study, multivariate analysis showed that the International Prognostic Index was the only independent predictor of FFS (relative risk, 2.3; Fig 7). The International Prognostic Index, or modifications thereof, along with the Berard method of grading, should be useful in the design of future therapeutic trials of patients with FL and in the selection of appropriate therapeutic approaches for individual patients.

REFERENCES


42. Kennedy JC, El-Badawy N, DeRose PB, Cohen C: Comparison of cell proliferation in breast carcinoma using image analysis (Ki-67) and flow cytometric systems. Anal Quant Cytol Histol 14:304, 1992


Prognostic value of cellular proliferation and histologic grade in follicular lymphoma

AR Martin, DD Weisenburger, WC Chan, EI Ruby, JR Anderson, JM Vose, PJ Bierman, MA Bast, DT Daley and JO Armitage

Updated information and services can be found at:
http://www.bloodjournal.org/content/85/12/3671.full.html
Articles on similar topics can be found in the following Blood collections

Information about reproducing this article in parts or in its entirety may be found online at:
http://www.bloodjournal.org/site/misc/rights.xhtml#repub_requests

Information about ordering reprints may be found online at:
http://www.bloodjournal.org/site/misc/rights.xhtml#reprints

Information about subscriptions and ASH membership may be found online at:
http://www.bloodjournal.org/site/subscriptions/index.xhtml