A High Tumor-Associated Macrophage Content Predicts Favorable Outcome in Follicular Lymphoma Patients Treated with Rituximab and Cyclophosphamide-Doxorubicin-Vincristine-Prednisone

Minna Taskinen,1,2 Marja-Liisa Karjalainen-Lindsberg,3 Heidi Nyman,1,2 Leena-Maija Eerola,1,2 and Sirpa Leppä1,2

Abstract

Purpose: Tumor-associated macrophage (TAM) content predicts survival in follicular lymphoma (FL) patients treated with chemotherapy. The aim of this study was to determine how combination of rituximab with chemotherapy influences TAM-associated clinical outcome.

Experimental Design: Expression of a macrophage marker, CD68, was determined immunohistochemically from FL samples of 96 patients treated with rituximab and cyclophosphamide-Adriamycin-vincristine-prednisone regimen. Of them, 71 received therapy at diagnosis and 25 at relapse. Neutrophil and CD3+ lymphocyte counts were also measured. The median follow-up time for the cohort was 54 months. Forty-five patients previously treated with chemotherapy served as a control group.

Results: Consistent with previous studies, high TAM amount was associated with adverse outcome in chemotherapy-treated patients (P = 0.026). In contrast, after rituximab and cyclophosphamide-doxorubicin-vincristine-prednisone regimen, high TAM content correlated with longer survival rates. According to Kaplan Meier estimates, the median progression free survival (PFS) was not reached for patients with high TAM content compared with 45 months for patients with low TAM scores (P = 0.006). A trend toward a better overall survival (OS) at 5 years was also observed for patients with high TAM content (OS, 97% versus 90%, P = 0.116). The positive prognostic value of TAMs was seen both for the patients treated at diagnosis and at relapse. In multivariate analyses, TAM content remained an independent prognostic factor for OS and PFS. Neutrophil and CD3+ lymphocyte counts did not correlate with outcome.

Conclusions: The data suggest that high TAM score is associated with a favorable prognosis in FL patients treated with immunochemotherapy.

Follicular lymphoma (FL) is the second most common subtype of all non–Hodgkin lymphomas accounting for ~20% of all adult cases in the western world. FL arises from B cells in the germinal center of lymphoid tissue. It is an indolent and a chemotherapy-sensitive disease, which is, however, considered rarely curable due to its propensity to relapse. Recently, however, a significant improvement of the outcome of patients has been obtained by combining a monoclonal anti-CD20 antibody, rituximab, with induction chemotherapy (1–3), or by prolonging the remission with rituximab maintenance therapy (4, 5). Despite the advances, response to treatment varies substantially among individual patients and outcome is often unpredictable. Treatment is also costly. These facts raise the need to identify more accurately the patients who benefit from immunochemotherapy.

In FL, a specific Follicular Lymphoma International Prognostic Index (FLIPI) has been proposed (6). Although FLIPI was developed before rituximab was established in the treatment of FL, it was recently shown to be a useful predictor of the outcome also in response to rituximab and cyclophosphamide-Adriamycin-vincristine-prednisone (R-CHOP) regimen (7). In estimating the prognosis of FL, FLIPI seems to distribute the patients into risk groups better than the International Prognostic Index for aggressive non–Hodgkin lymphomas (8) and is likely to be a helpful tool in selecting the most appropriate treatment for individual FL patients. However, five clinical characteristics of FLIPI do not provide information on the biological and molecular features of the FL.

Differences in the tumor microenvironment have been shown to associate with clinical course and outcome of FL patients treated with chemotherapy (9). In particular, tumor-infiltrating cytotoxic and regulatory T cells and macrophages have prognostic effect in FL (10–16). However, all biological
Data have been produced before rituximab has been adapted into clinical use, and there are currently no biomarkers that have been revalidated and shown to be prognostic in patients treated with rituximab-containing regimens. Our recent microarray and immunohistochemical pilot study suggests that gene expression by nonmalignant tumor cells has prognostic effect also in R-CHOP–treated FL patients (17). Taking into account these previous observations, we evaluated the predictive value of tumor-infiltrating inflammatory cells and especially macrophages in the era of R-CHOP treatment.

### Materials and Methods

**Patients.** A total of 141 FL patients treated at the Helsinki University Central Hospital were included in the study. Of these, 96 FL patients received R-CHOP in front-line (n = 71) or relapse (n = 25) (post-rituximab era), whereas 45 patients treated with chemotherapy or radiotherapy before rituximab was adapted into clinical routine served as a control group (pre-rituximab era). All patients in the post-rituximab group received rituximab for the first time. The relapsed patients had received various treatments front-line, including chlorambucil, combination chemotherapy, and local irradiation. All tissue samples were taken before treatment. Lymphoma classifications, including histopathology and immunophenotyping, were done at the Department of Pathology at Helsinki University Central Hospital Laboratoires according to the WHO classification. Treatment records of all patients were reviewed to confirm the appropriate treatment protocols and to document clinical characteristics, prognostic factors, and long-term follow-up. The protocol and sampling were approved by Institutional Review Board and Finnish National Authority for Medicolegal Affairs.

**Immunohistochemistry.** Immunohistochemistry was done on formalin-fixed, paraffin-embedded tissue sections either on individual slides or as a part of a tissue microarray (TMA) as described earlier (17). The sections were stained with anti-CD68 (1:2,000; clone KP1, Dako), anti-CD3 (1:100; Novocastra Laboratories Ltd.), or with Leder stain (18) to detect macrophages, T lymphocytes, and neutrophils, respectively.

The number of CD68+ cells was counted as absolute cell numbers at a 630× magnification with a Leica DM LB bright-field microscope (Leica Microsystems GmbH). With TMA, two representative fields with the most abundant immune infiltration/core were counted, resulting in

### Table 1. Characteristics of R-CHOP–treated and control patients (pre-rituximab group) with FL according to TAM content

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>R-CHOP group</th>
<th>Control group (pre-rituximab group)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All patients, N = 96 (%)</td>
<td>Low TAM content (&lt;67%), n = 64 (%)</td>
</tr>
<tr>
<td>Gender Female</td>
<td>58 (60)</td>
<td>37 (58)</td>
</tr>
<tr>
<td>Male</td>
<td>38 (40)</td>
<td>27 (42)</td>
</tr>
<tr>
<td>Age &lt;60</td>
<td>65 (68)</td>
<td>46 (72)</td>
</tr>
<tr>
<td>≥60</td>
<td>31 (32)</td>
<td>18 (28)</td>
</tr>
<tr>
<td>State of disease</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>71 (74)</td>
<td>48 (75)</td>
</tr>
<tr>
<td>Relapse</td>
<td>25 (26)</td>
<td>16 (25)</td>
</tr>
<tr>
<td>1. Relapse</td>
<td>19</td>
<td>13</td>
</tr>
<tr>
<td>2. Relapse</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Grade I</td>
<td>35 (36)</td>
<td>29 (45)</td>
</tr>
<tr>
<td>II</td>
<td>21 (22)</td>
<td>12 (19)</td>
</tr>
<tr>
<td>III</td>
<td>6 (6)</td>
<td>2 (3)</td>
</tr>
<tr>
<td>Missing</td>
<td>34 (35)</td>
<td>21 (33)</td>
</tr>
<tr>
<td>FLIPI 0-2</td>
<td>67 (70)</td>
<td>44 (69)</td>
</tr>
<tr>
<td>3-5</td>
<td>22 (23)</td>
<td>15 (23)</td>
</tr>
<tr>
<td>Missing</td>
<td>7 (7)</td>
<td>5 (8)</td>
</tr>
<tr>
<td>R-FLIPI 0-2</td>
<td>59 (61)</td>
<td>39 (61)</td>
</tr>
<tr>
<td>3-5</td>
<td>30 (31)</td>
<td>21 (33)</td>
</tr>
<tr>
<td>Missing</td>
<td>7 (7)</td>
<td>4 (6)</td>
</tr>
</tbody>
</table>

Abbreviation: NA, not applicable.

### Table 2. Cox’s proportional hazard regression analysis for R-CHOP-treated patients

<table>
<thead>
<tr>
<th>Factor</th>
<th>RR (95% CI)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>OS FLIPI</td>
<td>2.386 (1.224-4.651)</td>
<td>0.012</td>
</tr>
<tr>
<td>TAM score</td>
<td>0.907 (0.835-0.986)</td>
<td>0.023</td>
</tr>
<tr>
<td>PFS R-FLIPI</td>
<td>1.477 (1.097-1.989)</td>
<td>0.010</td>
</tr>
<tr>
<td>TAM score</td>
<td>0.972 (0.943-1.001)</td>
<td>0.055</td>
</tr>
</tbody>
</table>

NOTE: For FLIPI, the higher the number means worse. TAM content was included in the analysis as a continuous variable. Low TAM score is worse than high.

Abbreviations: RR, relative risk; CI, confidence interval.
two to six areas (one to three replicas). With sections on individual slides, three to four representative areas were counted, depending on the size of a section. Scoring results were averaged, and the consistency between TMA replicas and different scoring areas was evaluated and considered necessary for the sample to be included in the study. Because of the variation in sample size (greater possibility to include hotspots in whole tissue sections), cutoff points were analyzed separately in whole tissue sections and TMA cores. To confirm the correlation between the data from TMA and whole tissue sections, a subset of 20 samples was analyzed for CD68 immunoreactivity from both series. In these cohorts, the median TAM contents for TMA and whole tissue slides per high power fields were 21 (range, 9-44) and 35 (range, 22-55), respectively. The values for the highest tertiles were 28 and 40. A comparison between two series showed a significant correlation (\( P = 0.020 \)) with a correlation coefficient of 0.516. CD3+ cell counts were semiquantitatively assessed as 0 to 10%, 10% to 25%, 25% to 50%, or >50% relative to total number of lymphocytes. For neutrophils, the number of positive cells was counted from two to three fields per sample with 400× magnification.

### Statistics

To evaluate the correlation between the data from TAM and whole tissue sections, Pearson correlation coefficient was calculated. The \( \chi^2 \) test and Mann-Whitney nonparametric \( U \) test were used to assess differences in the frequency of prognostic factors. Survival rates were estimated by the Kaplan-Meier method and the differences between the subgroups were compared by the log-rank test. Overall survival (OS) was measured from the date of diagnosis until the last follow up or death from any cause. Progression free survival (PFS) was determined as an interval between the first day of therapy and the date of relapse or death. Both univariate and multivariate analyses on the prognostic effect of identified factors were done using Cox proportional hazards model. The factors were first included in the univariate analyses. The predictor variables, which had statistically significant effect on survival, were included in multivariate analyses. TAM and neutrophil contents were analyzed as continuous variables, whereas CD3+ cell counts were categorized into four subgroups. Probability values below 0.05 were considered statistically significant. All \( P \) values were two-tailed. Data were analyzed using SPSS 11.0 (SPSS, Inc.).

### Results

#### Patient and disease characteristics.

The baseline characteristics of the 96 R-CHOP–treated patients are listed in Table 1. The median age of the cohort was 55 years (range, 27-79 years). The median follow up from diagnosis was 54 months for all patients, 46 months for 71 patients treated front-line, and 114 months for 25 patients treated in relapse. In total, 11 patients had died at the time of last follow up. The predicted 5-year PFS and OS for all patients were 47% and 92%, respectively. The patients were further divided into two subgroups based on TAM content. The median levels for TAMs per high power field and square millimeter were 21 (range, 9-44) and 222 (range, 0-812), respectively. The corresponding numbers for the highest tertiles were 30 and 310. When the highest tertile was...
used as a cutoff value between two groups (TAM-low and TAM-high), no differences were observed in the frequency of baseline characteristics between the groups (Table 1).

**Survival analyses.** Cox univariate analysis was done to assess the prognostic value of tumor-infiltrating inflammatory cells on OS. TAM and neutrophil contents were analyzed as continuous variables, whereas CD3+ cell counts were categorized into four subgroups. Besides FLIPI at diagnosis ($P = 0.033$) and before R-CHOP (R-FLIPI, $P = 0.021$), only significant prognostic factor for OS was TAM score ($P = 0.030$). Neutrophil and CD3+ lymphocyte scores had no significant influence on OS (for both, $P = \text{nonsignificant}$). When TAM score was included in the multivariate analyses with FLIPI or R-FLIPI, both factors had independent prognostic value for OS and PFS, respectively (Table 2).

In Kaplan-Meier analyses, the cutoff level of 67%, rather than 50% or 90%, was found to better discriminate between subgroups with different outcomes (Table 3). The 5-year OS for R-CHOP–treated patients with lower TAM score (<67%) was 90% compared with 97% of those with higher TAM levels (>67%, $P = 0.116$; Fig. 1A; Table 3). Likewise, the 5-year OS was 98% for the low FLIPI scores (0-2) at diagnosis and 86% for the high scores (3-5; $P = 0.068$).

For all patients, PFS rates after R-CHOP regimen were significantly better among the patients with high than low TAM scores (median not reached versus 45 months, $P = 0.006$; Fig. 1B). In comparison, R-FLIPI score identified before R-CHOP therapy could separate the low- and intermediate-risk patients from high-risk groups (median not reached versus 35 months, $P = 0.017$; Fig. 1C).

To determine how the treatments given at different disease stages contributed to OS, PFS rates were estimated separately after front-line and second-line therapies. PFS rates from the date of first front-line R-CHOP tended to be better for the patients with high TAM content than for the ones with low scores (median PFS not reached versus 48 months, $P = 0.067$; Fig. 2A). When the outcome of 25 relapsed patients were compared, a trend toward better PFS was also seen in patients with high TAM content (median PFS, 57 months versus 22 months, $P = 0.075$; Fig. 2B).

Finally, we evaluated the prognostic effect of TAM scores on the outcome of small group of 45 patients who received initial therapy before rituximab was available in clinical routine (pre-rituximab era). The baseline characteristics for the pre-rituximab group are shown in Table 1. In comparison with R-CHOP-treated patients, the pre-rituximab group contained less high-risk patients according to FLIPI at diagnosis ($P = 0.003$). For all other comparisons, including R-FLIPI scores, no differences were observed ($P > 0.14$ for all of the analyses). The median levels for TAMs per high power field and square millimeter were 21 (range, 0-62) and 222 (range, 0-650), respectively. The corresponding numbers for the highest tertiles were 31 and 323. No differences were observed in TAM levels between pre-rituximab and R-CHOP groups ($P = 0.72$). When the 90% was used as a cutoff point, PFS was found to be worse in patients with high TAM score compared with low scores (median, 17 months versus 41 months; $P = 0.026$). The result confirms previous findings on chemotherapy-treated patients (13). When the same patient cohort received rituximab-containing regimen at relapse, high TAM score (>67%) was associated with better PFS (median, 57 months versus 17 months; $P = 0.009$). Together, the data suggest that addition of rituximab to chemotherapy reverses the negative prognostic effect of high TAM content to favorable.

**Discussion**

In the present study, we show that TAM content in the diagnostic lymph node is a significant independent prognostic...
factor in FL patients treated with R-CHOP regimen. Higher TAM scores seem to protect the patients from early relapse and identify a subgroup of patients with favorable prognosis.

According to gene expression profiling studies in FL, the composition of tumor microenvironment has been defined as an important determinant of survival (9, 15, 17, 19). The leukemia/lymphoma molecular profiling project determined two major signatures with significantly different outcomes (9). One of these signatures, called immune response-2, included high expression of genes, encoding proteins in follicular dendritic cells and macrophages, and was associated with inferior outcome. Subsequent studies have analyzed the relation between the outcome and macrophages with immuno-histochimistry. Thus far, however, the results on the prognostic effect of CD68-positive macrophages in the diagnostic FL tissue have been inconsistent. Whereas Farinha et al. (13) showed that high TAM score is associated with adverse outcome in response to chemotherapy, other studies could not observe prognostic effect of TAMs for outcome (11, 14). The discrepancies could be related to different patient populations and technical factors, such as quantification methods and different cutoff levels used to distinguish the high and low subgroups.

Using similar scoring method and cutoff level as Farinha et al. (90%; ref. 13), we could reproduce the association of high TAM score with unfavorable PFS among the patients not receiving rituximab in part of their initial therapy. However, our major interest was to determine whether TAM content has prognostic effect for FL patients treated with combination of rituximab and chemotherapy. The finding that high TAM content turned out to predict favorable PFS in response to combination of rituximab and chemotherapy at relapse in the patient cohort who received chemotherapy in the first line illustrates that addition of rituximab to chemotherapy has a strong influence on the prognostic effect of TAMs in FL. This, together with the data on the larger cohort of patients treated homogenously with R-CHOP regimen, encourages us to believe that addition of rituximab to chemotherapy reverses the negative prognostic effect of high TAM content to favorable. The data also highlight the importance to revaluate all previously identified prognostic factors in the post-rituximab era.

In our material, TAM content predicted outcome both as a continuous and stratified variable. The finding is important because the cutoff levels best discriminating the low and high subgroups were different between chemotherapy- and immuno-chemotherapy-treated patients. The frequency of macrophages in the diagnostic FL tissue also differs in reported series. For example, the relative number of TAMs in our cohort was higher than that reported previously by Farinha et al. (13). Considering that younger patients tended to have somewhat lower TAM levels in our patients cohort and the fact that the median age of the patients was only 44 years in the previous study, the differences may be partially age related. However, it seems obvious that different patient cohorts have many intrinsic properties, which account for outcomes. Thus, comparison of data between different series should be done with extreme caution. It is also important to emphasize that we do not consider certain cutoff points to be biologically significant, but rather think that TAM levels form a continuum with increased numbers of TAMs correlating with improved survival in response to R-CHOP regimen.

The mechanism by which the addition of rituximab to chemotherapy improves the outcome of the patients with high TAM content is unknown but is likely to be related to antibody-mediated cellular cytotoxicity. The antitumoral activity of rituximab is dependent on Fc receptor–mediated interactions with effector cells, including neutrophils, natural killer cells, and macrophages (20, 21). Considering an association between the Fc receptor IIIa genotype and clinical and molecular responses to rituximab (21), Fc receptor IIIa polymorphism may contribute macrophage-dependent antibody–mediated cellular cytotoxicity activity against lymphoma cells. Macrophages can eliminate B lymphocytes by direct Fc receptor–mediated phagocytosis. Alternatively, they may secrete cytolytic factors or release cytokines, thereby recruiting other effector cells to amplify the inflammatory response. Emphasizing that tissue macrophages are critical for B-cell depletion after anti-CD20 therapy (22, 23), it is possible that there is an interrelationship between TAM content and efficacy of rituximab.

In conclusion, we have shown that high TAM content predicts the outcome of FL patients in response to R-CHOP regimen. Our data not only confirm that nonmalignant tumor cells have a profound prognostic effect in FL but also illustrate that molecular prognostic factors need to be reevaluated in the post-rituximab era.

Acknowledgments
We thank Anneli Grönqvist and Onerva Levalampi for technical assistance and Molecular Imaging Unit at the University of Helsinki for providing instruments and support with microscopy.

References


A High Tumor-Associated Macrophage Content Predicts Favorable Outcome in Follicular Lymphoma Patients Treated with Rituximab and Cyclophosphamide-Doxorubicin-Vincristine-Prednisone


Updated version
Access the most recent version of this article at:
http://clincancerres.aacrjournals.org/content/13/19/5784

Cited articles
This article cites 23 articles, 18 of which you can access for free at:
http://clincancerres.aacrjournals.org/content/13/19/5784.full.html#ref-list-1

Citing articles
This article has been cited by 38 HighWire-hosted articles. Access the articles at:
/content/13/19/5784.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.