Impact of histological grading on survival in the SWOG S0016 follicular lymphoma cohort

This study examined the impact of higher grade follicular lymphoma on early progression and survival as well as the reproducibility of grading in a large prospective, phase III randomized intergroup trial, S0016, comparing six cycles of cyclophosphamide, doxorubicin, vincristine, prednisone (CHOP) plus rituximab with six cycles of CHOP followed by iodine I-131 tositumomab radioimmunotherapy. We show a lack of correlation between histological grade and patients' outcome in the rituximab plus doxorubicin-containing chemotherapy era; and that grade is often revised, and lowered, during expert hematopathology review. Therefore, other more reproducible biomarkers of early progression should be sought for FL.

FL is a complex disease with variable histological appearances and clinical outcomes. There are clinical and genetic indices with prognostic capability, including the FL International Prognostic Index (FLIPI) and the more recent M7-FLIPI, which incorporates the mutation status of seven genes with superior power to distinguish patients' risk.[1,2] In addition to baseline prognostic tools, early progression of disease at 24 months following chemotherapy with CHOP plus rituximab is a powerful predictor of overall survival.[3] The original histological grading scheme was developed prior to these various methods and is still retained in the revisions to the World Health Organization lymphoma classification scheme.[4,5]

According to the current World Health Organization criteria, FL is divided into grades based on the average number of centroblasts from ten representative high power fields, taking into account areas with both few and frequent centroblasts. Grades 1 and 2 are defined by an average of 0-5 and 6-15 centroblasts, respectively, per high power field and are now considered together as low grade FL (hereafter referred to as FL1/2). Grade 3, with more than 15 centroblasts per high power field, is further subdivided into 3A with admixed centrocytes and 3B without admixed centrocytes with the latter representing an essentially pure population of centroblasts. Using current criteria, FL3A and FL3B must demonstrate a follicular pattern (otherwise the diagnosis is moved into the diffuse large B-cell lymphoma category). Although centroblast counts have been a staple of FL grading since the 1980s, the reproducibility of this method is debatable and the methods somewhat complex and time-consuming.[6,7] Furthermore, there are conflicting reports about whether there is a significant biological or survival difference between grades 3A and 3B. Some reports note that survival is similar while others studies found that FL3B has unique biological features.[8,9] These conflicting data challenge the clinical utility of grading distinctions with current therapeutic options.

The importance of FL grade has not been examined in a large uniformly treated cohort, particularly using progression of disease within 2 years after diagnosis, which is strongly associated with poor outcomes in patients receiving first-line treatment with CHOP plus rituximab.[3] The South West Oncology Group (SWOG) cooperative research consortium previously conducted a phase III randomized intergroup trial, S0016, comparing six cycles of CHOP plus rituximab with six cycles of CHOP followed by iodine I-131 tositumomab radioimmunotherapy. The study initially showed no significant improvement in progression-free survival between the two arms (although both progression-free survival and overall survival were outstanding on both arms).[10,11] The recent 10-year update of the trial demonstrated a progression-free survival benefit, but not overall survival benefit in the radioimmunotherapy arm.[12] A subsequent S0016 analysis evaluated clinical factors and found that lactate dehydrogenase, β_2 microglobulin, and FLIPI scores were significantly correlated with outcome, but did not address grade or histological pattern.[10] S0016 is an excellent cohort for assessment of the prognostic significance of FL grading and pattern due to the use of immunochemotherapy, uniform treatment and outcome data, and centralized pathology review. We therefore sought to evaluate whether FL grading continues to be an important clinical parameter in the modern treatment era.

We previously confirmed the diagnoses of FL for S0016 at the time of study closure.[11] This is a mature cohort with a median follow up of 10.2 years (range, 50 days to 15.3 years). Briefly, the available slides and associated pathology reports were reviewed by one of three SWOG Lymphoma Committee pathologists (RMB, CMS, and LMR) for eligibility according to the recently revised World Health Organization criteria.[5] Given the known, high reproducibil-

Table 1. Survival stratified by histological grade.

		Patients	Patients CHOP-R	CHOP-RIT	
Total		491 (100%)	246 (100.0%)	245 (100%)	
FL 1/2		452 (92%)	227 (92.2%)	225 (91.8%)	
FL 3A		39 (7.9%)	19 (7.7%)	20 (8.1%)	
Grade	**Total**	**Failed**	**Progression-Free Survival Estimate of 10-years PFS**	**95% CI**	***P***
FL 1/2	452	227	49%	44.5%-54.1%	
FL 3A	39	20	47%	30.1%-61.7%	0.789
Grade	**Total**	**Failed**	**Overall Survival Estimate of 10-years OS**	**95% CI**	***P***
FL 1/2	452	95	79%	(74.6%-82.6%)	
FL 3A	39	12	68%	(50.3%-80.5%)	0.0990

CHOP: cyclophosphamide, doxorubicin, vincristine, prednisone; R: rituximab; RIT: radioimmunotherapy; FL: follicular lymphoma; CI: confidence interval; OS: overall survival; PFS: progression free survival.

ity of diagnosis of FL1/2, if the reviewing SWOG pathologist agreed with the original diagnosis of FL1/2, no further review was conducted. If the SWOG pathologist changed the diagnosis or changed the grade, the case was brought to "consensus" review, during which all three SWOG pathologists reviewed the case at a multi-headed microscope, discussed the findings, and agreed on the final diagnosis and grade. In addition, because of the known low reproducibility of FL3, all FL3A and FL3B cases were confirmed by the panel.[6]

For the survival analyses, cases with grade 1/2 were considered together as low grade; cases of pure grade 3A were considered 3A; cases with mixed low and high grades (combinations of grade 1/2 with grade 3A) were included in the 3A category. A single case of pure FL3B was excluded given its unicity. Cases with mixed FL and diffuse large B-cell lymphoma were considered as diffuse large B-cell lymphoma and were not included in this FL survival analysis. Survival was assessed using the methods of Kaplan and Meier.[13] We estimated 10-year progression-free and overall survival rates and compared the results with a log-rank test.

The SWOG pathology reviewers' consensus diagnoses were used for the survival analyses. A total of 491 patients were assessed for histological grade, of whom 452 (92.0%) had grade 1/2 and 39 (7.9%) patients had grade 3A. No cases were considered pediatric-type FL, since this trial was limited to adult patients with advanced stage disease. The distribution of FL1/2 or FL3A cases was similar between treatment arms (91.8% and 8.1% for CHOP-immunochemotherapy and 92.2% and 7.7% for CHOP plus rituximab, respectively). These results are summarized in Table 1.

The estimated 10-year overall survival for patients with FL1/2 was 79% [95% confidence interval (95% CI): 74.6-82.6%], while that for patients with for FL3A was 68% (95% CI: 50.3-80.5%) (log-rank *P*-value=0.0990). In addition, the estimated 10-year progression-free survival was 49% for patients with FL1/2 (95% CI: 44.5- 54.1%); and 47% for those with FL3A (95% CI: 30.1-61.7%) (log-rank *P*-value=0.7897). Survival curves are illustrated in Figure 1.

A total of 521 cases were submitted with a diagnosis of FL. Of these, 497 (95.4%) had a definitive grade in the diagnostic line or description from the submitted outside pathology report. Of the 497 graded FL cases, changes to that grade occurred in 54 (10.8%). Interestingly, 45 of these 54 (83.3%) changes were down grades from grade FL3A to FL1/2, while in contrast, only 9/54 (16.7%) were upgraded from an original diagnosis of FL1/2 to FL3A. Collectively, these data indicated that most FL grading changes by the SWOG review panel were decreases in the FL grade and highlighted a bias towards higher grades in community practice. Microscopic review revealed that the down-graded FL cases often had either large centrocytes, sclerosis that may distort cell cytology, or many follicular dendritic cells, which may have made them particularly challenging.

Both the impact of higher grade FL on early progression and survival as well as the reproducibility of grading in a large prospectively treated cohort were investigated in this study. We found that: (i) the impact of histological grading on patients' outcome in the rituximab plus doxorubicin-containing chemotherapy era is negligible; and (ii) there is a tendency to "over-grade" FL. In the USA, bendamustine has largely replaced CHOP as the preferred chemotherapy platform for newly diagnosed advanced stage FL, based upon randomized data suggesting similar to improved progression-free survival and a better toxicity profile.[14,15] These two randomized trials excluded patients with grade 3 FL. In agreement with Koch *et al.*,[16] our results suggest that, at least when doxorubicin-containing chemotherapy with an anti-CD20 monoclonal antibody is utilized, the outcome of patients with grade 3A FL is similar to that of patients with grade 1/2 disease. These observations should be validated in a cohort of patients treated with bendamustine plus rituximab. Similar to an earlier study, there were no cases showing combinations of both FL3A and FL3B, highlighting that the pathogenesis of these two diseases is likely different.[16]

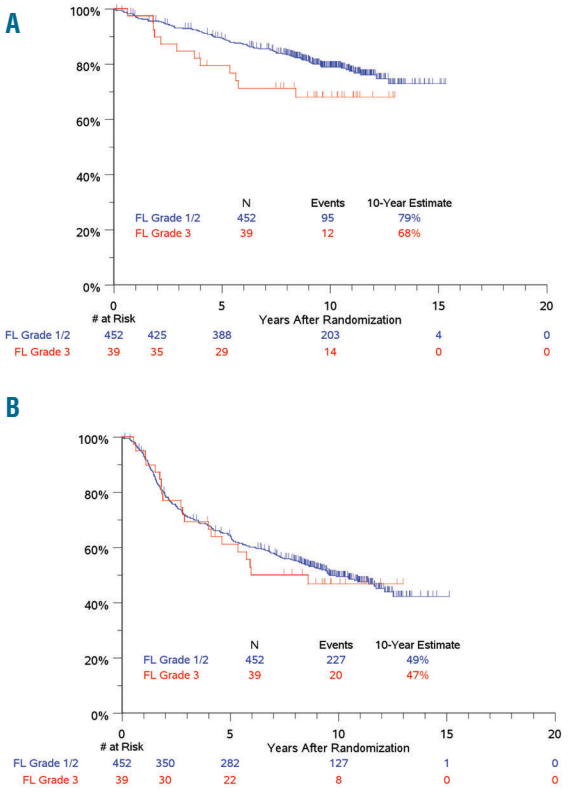

Figure 1. Survival according to histological grade. (A) Overall survival. (B) Progression free survival.

Challenges to the reproducibility in grading, previously reported in the literature, were again found in this study. The most frequent change was downgrading of originally denoted FL3A or FL3B cases to FL1/2. Our findings highlight a tendency to over-grade FL in routine practice most likely due to the difficulty in cytologically separating the various cell types within the germinal center environment. Various centroblast "mimics" such as large centrocytes (large cleaved cells but without nucleoli) or follicular dendritic cells (single or binucleated cells with clear nuclei and centrally placed eosinophilic nucleoli) can be difficult to identify and exclude. In agreement with our findings, grade 3 FL was one of the least reproducible of the diagnostic categories of lymphoma in the National Cancer Care Network database study.[6] As the FL field moves forward, integrated biomarkers such as clinical features with genomics and/or expression markers are likely to be more important than grade in prognosis, and may also provide predictive information for specific therapies.

Lisa M. Rimsza,[1] Hongli Li,[2] Rita M. Braziel,[3] Catherine M. Spier,[4] Daniel O. Persky,[5] Jennifer Dunlap,[3] Michael LeBlanc,[2] Nancy Bartlett,[6] John P. Leonard,[7] Sonali M. Smith,[8] Oliver W. Press[2] and Jonathan W. Friedberg[9]

[1]Department of Laboratory Medicine and Pathology, Mayo Clinic, Scottsdale, AZ; [2]Fred Hutchinson Cancer Research Center, Seattle, WA; [3]Department of Pathology, Oregon Health Sciences University, Portland, OR; [4]Department of Pathology, University of Arizona, Tucson, AZ; [5]University of Arizona Cancer Center, University of Arizona, Tucson, AZ; [6]Washington University-Siteman Cancer Center, St. Louis, MO; [7]Weill Cornell Medicine and New York Presbyterian Hospital, New York, NY; [8]Department of Medicine, University of Chicago, Chicago, IL and [9]James P. Wilmot Cancer Center, University of Rochester, Rochester, NY, USA

Funding: research reported in this publication was supported by the National Cancer Institute of the National Institutes of Health under Award Numbers U10CA180888, U10CA180819, U10CA180828 and U10CA180821; U10CA13612, U10CA11083 and in part by GlaxoSmithKline plc. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or GlaxoSmithKline plc.

Correspondence: rimsza.lisa@mayo.edu
doi:10.3324/haematol.2017.175059

Information on authorship, contributions, and financial & other disclosures was provided by the authors and is available with the online version of this article at www.haematologica.org.

References

1. Pastore A, Jurinovic V, Kridel R, et al. Integration of gene mutations in risk prognostication for patients receiving first-line immunochemotherapy for follicular lymphoma: a retrospective analysis of a prospective clinical trial and validation in a population-based registry. Lancet Oncol. 2015;16(9):1111-1122.
2. Solal-Celigny P, Roy P, Colombat P, et al. Follicular lymphoma international prognostic index. Blood. 2004;104(5):1258-1265.
3. Casulo C, Byrtek M, Dawson KL, et al. Early relapse of follicular lymphoma after rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone defines patients at high risk for death: an analysis from the National LymphoCare Study. J Clin Oncol. 2015;33(23):2516-2522.
4. Mann RB, Berard CW. Criteria for the cytologic subclassification of follicular lymphomas: a proposed alternative method. Hematol Oncol. 1983;1(2):187-192.
5. Swerdlow SH, Campo E, Pileri SA, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. Blood. 2016;127(20):2375-2390.
6. LaCasce AS, Kho ME, Friedberg JW, et al. Comparison of referring and final pathology for patients with non-Hodgkin's lymphoma in the National Comprehensive Cancer Network. J Clin Oncol. 2008;26(31):5107-5112.
7. Metter GE, Nathwani BN, Burke JS, et al. Morphological subclassification of follicular lymphoma: variability of diagnoses among hematopathologists, a collaborative study between the Repository Center and Pathology Panel for Lymphoma Clinical Studies. J Clin Oncol. 1985;3(1):25-38.
8. Ott G, Katzenberger T, Lohr A, et al. Cytomorphologic, immunohistochemical, and cytogenetic profiles of follicular lymphoma: 2 types of follicular lymphoma grade 3. Blood. 2002;99(10):3806-3812.
9. Wahlin BE, Yri OE, Kimby E, et al. Clinical significance of the WHO grades of follicular lymphoma in a population-based cohort of 505 patients with long follow-up times. Br J Haematol. 2012;156(2):225-233.
10. Press OW, Unger JM, Rimsza LM, et al. A comparative analysis of prognostic factor models for follicular lymphoma based on a phase III trial of CHOP-rituximab versus CHOP + 131iodine--tositumomab. Clin Cancer Res. 2013;19(23):6624-6632.
11. Press OW, Unger JM, Rimsza LM, et al. Phase III randomized intergroup trial of CHOP plus rituximab compared with CHOP chemotherapy plus (131)iodine- tositumomab for previously untreated follicular non-Hodgkin lymphoma: SWOG S0016. J Clin Oncol. 2013;31(3):314-320.
12. Shadman M, Li H, Rimsza LM, et al. Continued excellent outcomes in previously untreated follicular lymphoma patients after treatment with CHOP plus rituximab or CHOP plus (131) iodine-tositumomab - long term follow-up of phase II randomized study SWOG 0016. Blood. 2016;128(22):616.
13. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc. 1958;53(282):457-481.
14. Flinn IW, van der Jagt R, Kahl BS, et al. Randomized trial of bendamustine- rituximab or R-CHOP/R-CVP in first-line treatment of indolent NHL or MCL: the BRIGHT study. Blood. 2014;123(19):2944-2952.
15. Rummel MJ, Niederle N, Maschmeyer G, et al. Bendamustine plus rituximab versus CHOP plus rituximab as first-line treatment for patients with indolent and mantle-cell lymphomas: an open-label, multicentre, randomised, phase 3 non-inferiority trial. Lancet. 2013;381(9873):1203-1210.
16. Koch K, Hoster E, Ziepert M, et al. Clinical, pathological and genetic features of follicular lymphoma grade 3A: a joint analysis of the German low-grade and high-grade lymphoma study groups GLSG and DSHNHL. Ann Oncol. 2016;27(7):1323-1329.