

Revisiting the immune microenvironment of diffuse large B-cell lymphoma using a tissue microarray and immunohistochemistry: robust semi-automated analysis reveals CD3 and FoxP3 as potential predictors of response to R-CHOP

Rita Coutinho,¹ Andrew J. Clear,¹ Emanuele Mazzola,² Andrew Owen,^{1,3} Paul Greaves,¹ Andrew Wilson,¹ Janet Matthews,¹ Abigail Lee,^{1,3} Rute Alvarez,⁴ Maria Gomes da Silva,⁴ José Cabeçadas,⁵ Donna Neuberg,³ Maria Calaminici,^{1,3} and John G. Gribben¹

¹Department of Hemato-Oncology, Barts Cancer Institute, Queen Mary University, London, UK; ²Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, Boston, MA, USA; ³Department of Histopathology, Barts Health NHS Trust, Royal London Hospital, UK; ⁴Department of Hematology, Portuguese Institute of Oncology, Lisbon, Portugal; and ⁵Department of Pathology, Portuguese Institute of Oncology, Lisbon, Portugal

ABSTRACT

Gene expression studies have identified the microenvironment as a prognostic player in diffuse large B-cell lymphoma. However, there is a lack of simple immune biomarkers that can be applied in the clinical setting and could be helpful in stratifying patients. Immunohistochemistry has been used for this purpose but the results are inconsistent. We decided to reinvestigate the immune microenvironment and its impact using immunohistochemistry, with two systems of image analysis, in a large set of patients with diffuse large B-cell lymphoma. Diagnostic tissue from 309 patients was arrayed onto tissue microarrays. Results from 161 chemoimmunotherapy-treated patients were used for outcome prediction. Positive cells, percentage stained area and numbers of pixels/area were quantified and results were compared with the purpose of inferring consistency between the two semi-automated systems. Measurement cut-points were assessed using a recursive partitioning algorithm classifying results according to survival. Kaplan-Meier estimators and Fisher exact tests were evaluated to check for significant differences between measurement classes, and for dependence between pairs of measurements, respectively. Results were validated by multivariate analysis incorporating the International Prognostic Index. The concordance between the two systems of image analysis was surprisingly high, supporting their applicability for immunohistochemistry studies. Patients with a high density of CD3 and FoxP3 by both methods had a better outcome. Automated analysis should be the preferred method for immunohistochemistry studies. Following the use of two methods of semi-automated analysis we suggest that CD3 and FoxP3 play a role in predicting response to chemoimmunotherapy in diffuse large B-cell lymphoma.

Introduction

Incorporation of rituximab into standard treatment has clearly improved the outcome of patients with diffuse large B-cell lymphoma (DLBCL),^{1,2} although patients with primary refractory disease or relapse have emerged as a particularly difficult group to cure. Indeed, there is an urgent need for novel therapeutic approaches in these patients.

The role of the microenvironment in DLBCL biology and outcome gained relevance when independent gene expression profiling defined distinct biological traits that were driven by the non-malignant cells in the tumors.³⁻⁶ However, gene expression profiling data require further validation and need to be made simpler in order to be useful for clinical trial design and for clinical practice.

Immunohistochemistry (IHC) has been explored to enumerate and functionally characterize the microenvironment in DLBCL.⁷⁻¹³ IHC can be extended to clinical practice, which makes it highly attractive as a diagnostic and prognostic tool. Nevertheless, the results published regarding the immune microenvironment in DLBCL are contradictory. The use of

inconsistent methodology likely explains these results. Moreover, it is known that it is difficult to obtain reproducible results when counting cells across large tumor areas manually. Categorization of the density of cell infiltration is used to overcome this problem, but results might be misleading and there is a lack of validation of cutpoints.

The main scope of our study was to re-investigate the immune microenvironment of diagnostic samples from 309 patients with DLBCL by two different methods of semi-automated image analysis. These approaches are being used to analyze large areas of tumor, making them ideal for IHC studies exploring prognosis. We fully expected to detect a high degree of inconsistency between the results of the two systems, similar to that reported when manual and automated analyses are compared. As a secondary aim, we considered the prognostic role of different immune biomarkers in 161 patients treated with rituximab plus cyclophosphamide, doxorubicin, vincristine and prednisolone (R-CHOP). Total T-lymphocytes, their functional subsets and macrophages were studied.

We added robustness to our data by using two different

©2015 Ferrata Storti Foundation. This is an open-access paper. doi:10.3324/haematol.2014.110189

The online version of this article has a Supplementary Appendix.

Manuscript received on August 3, 2014. Manuscript accepted November 24, 2014.

Correspondence: rittakoutinho@gmail.com

systems of image analysis and demonstrate that, against our expectations, the agreement of the two systems with each other is strong, even when comparing different variables analyzed for each biomarker, such as cell density and percentage of area stained. By recursive partitioning we then identified two subsets of patients, with different infiltration of CD3 and FoxP3, which had different outcomes after R-CHOP treatment. Importantly, although different cutpoints were retrieved for each variable analyzed by the two different systems, the agreement in allocating a patient into a high or low density cohort was extremely good, suggesting that the prognostic value of these biomarkers established in this cohort is connected to underlying biological differences related to the immune microenvironment.

Methods

Patients' characteristics

Ethical approval for this study was obtained from Local Regional Ethics Boards. We identified 225 patients with *de novo* DLBCL (77 treated with R-CHOP) diagnosed at St. Bartholomew's Hospital with available clinical data and formalin-fixed, paraffin-embedded diagnostic biopsies suitable for the study. A further 84 R-CHOP-treated patients from the Portuguese Institute of Oncology (Lisbon) were evaluated. The cutpoint and outcome analyses were based on the R-CHOP dataset. The clinical characteristics and outcome of these patients are detailed in *Online Supplementary Table S1*. The median follow-up of the patients was 51 months.

Tissue microarray and immunohistochemistry

Tissue cores were taken from representative tumor regions identified on sections stained with hematoxylin and eosin and confirmed by CD20 staining. Information on the staining protocol and antibody panels is provided in the *Online Supplementary Appendix* and *Online Supplementary Table S2*.

Semi-automated image analysis

Ariol system

Slides were scanned using an Olympus BX61 microscope with an automated platform. Representative regions were selected for training, after which the number of positive cells/area of representative lymphoma tissue and the percentage stained area/area of representative lymphoma tissue were quantified for all antibodies. A mean value was estimated for each patient.

Pannoramic viewer system

Slides were scanned using the Pannoramic 250 Flash II scanner (3DHISTECH, Hungary). The number of stained pixels/area was quantified using the DensitoQuant module. A mean value was estimated for each patient.

Cutpoint determination

Cutpoint discrimination was assessed using the recursive splitting algorithm in the *rpart* package (<http://cran.r-project.org/web/packages/rpart/index.html>) within the R statistical software.¹⁴

To achieve a simple subdivision into two classes, the algorithm was forced to obtain a single split for each considered measurement, with a splitting criterion based on the Gini index.

Two to ten cross validations were performed for each measurement to check for the consistency of each cutpoint, showing no modifications in the outcomes.

Statistical analysis

Differences between the groups of patients were tested using the χ^2 or Fisher exact test, when appropriate. Association between pairs of biomarkers was tested using the Fisher exact test; for robustness, we only considered splits by the recursive algorithm that yielded ≥ 30 individuals in the smallest group. Pearson correlation was calculated for all IHC variables and clinical parameters.

The outcomes measured were overall survival, defined as the time from diagnosis to death from any cause, with surviving patients censored at last follow-up, and progression-free survival, defined as the time from diagnosis to failure of treatment (including not achieving complete response or relapse after complete response) or death from any cause.

For every quantified measurement survival was estimated using Kaplan-Meier estimators, and differences in the same measurement between groups were assessed using the log-rank test. To accommodate for the optimization method within the splitting algorithm we considered only log-rank *P* values < 0.01 as statistically significant. Multivariate analysis was performed using a Cox proportional hazards model (stepwise backwards and forwards methods) including IHC parameters and clinical factors in the International Prognostic Index with prognostic significance in univariate analysis. *P* values < 0.05 are considered statistically significant. Statistical analyses were performed using SPSS version 19.0 (SPSS), Prism version 5.03 (GraphPad Software) and R version 3.0.2.

Results

Heterogeneous density of immune cells in the microenvironment of patients with diffuse large B-cell lymphoma

The density of the studied biomarkers within a given tumor was homogeneous, as expected from the diffuse histological pattern of DLBCL. This validates interpretation in partially represented cores in which viable tissue is properly represented.

As expected, patients with DLBCL had heterogeneous infiltration of total T cells, their subsets and macrophages at diagnosis (*Online Supplementary Table S3*).

CD3⁺ T cells were the most abundant cell type studied. The median CD4:CD8 ratio for all cases analyzed was between 0.36 and 0.68.

FoxP3 is a CD4⁺ T-cell transcription factor expressed by regulatory T cells. Its expression is clear and discrete and therefore training and analysis were facilitated. The number of cells expressing FoxP3 varied between samples (median 228.6; range, 0 - 3197), but less so than other markers (e.g. TIA1, median 1648; range, 31.7 - 5859).

TIA1 is a cytoplasmic marker expressed by cytotoxic T cells, independently of their activation status, and natural killer cells. CD56 staining was done to explore the extent of infiltration by natural killer cells and was almost absent in our series (*data not shown*). The absence of natural killer cells suggests that TIA1 expression is specific to T cells. CD68 is a cytoplasmic lysosomal protein expressed in all tissue macrophages. Large and interdigitating cells, such as macrophages, are difficult to quantify.¹⁵ The image analysis training was challenging and focused on capturing the areas of dark brown IHC staining present in the cell bodies. CD68 staining was heterogeneous in this cohort.

Excellent concordance between two systems of automated image analysis for characterization of the microenvironment

Multiple studies have shown that manual quantification of abundantly expressed IHC markers is inconsistent. To overcome this problem a number of semi-automated systems for IHC analysis have been launched that allow analysis of large areas of staining and increase the consistency of results. However, each of these systems has its own features and scripts for analysis and it is unknown how robust results are when obtained from different automated systems.

We therefore expected to detect considerable inconsistency between the results of two analysis systems that use different methods of counting. This would have implications for the future of IHC studies. To test this hypothesis we used the Ariol and the Panoramic Viewer semi-automated systems to examine the immune microenvironment in DLBCL. The methodology is presented in detail in the *Online Supplementary Appendix* and *Online Supplementary Figure S1*.

We first compared two different measures obtained from the Ariol system: (i) number of positive stained cells/area of viable tumor and (ii) percentage stained area for each marker. As can be appreciated from Figure 1A-C and *Online Supplementary Figure S2*, we found an excellent correlation between the two Ariol measures for each of the markers studied (Pearson r -values between 0.92 and 0.98, $P < 0.0001$). Importantly the correlation was high even for the cytoplasmic markers TIA1 and CD68, for which training was more challenging and low inter-

observer agreements have been reported. These results suggest that, applying our methodology, any of the two measures retrieved by the Ariol system can be used for future IHC studies.

We then compared the Ariol results with those obtained with the Panoramic Viewer, where quantification of the DAB stained pixels/area is performed. With the Panoramic Viewer the tumor density is not taken into account and a precise selection of representative tumor areas is more demanding. Even so, we surprisingly detected significant correlations between the data retrieved from the two systems (Figure 1D-F). Correlations were excellent for T-cell markers (Pearson r -values between 0.89 and 0.92) and more modest for cytoplasmic proteins (CD68, $r = 0.77$; TIA1, $r = 0.67$). This indicates that, unlike manual analysis of IHC, semi-automated systems are robust and preferable for future IHC studies.

CD3 and FoxP3 are potential predictors of progression-free survival after R-CHOP therapy

Our secondary aim was to investigate the prognostic impact of immune biomarkers in DLBCL. For this purpose we specifically selected the 161 R-CHOP-treated patients. *Online Supplementary Table S4* details survival, hazard ratios, confidence intervals and P -values for all significant variables in univariate and multivariate analyses.

We did not detect any interactions between the clinical variables detailed in *Online Supplementary Table S1* and the biomarkers studied (*data not shown*). We applied strict quality criteria to the image analysis to ensure robustness and reproducibility of the techniques. This meant that images were rejected for a number of markers that did not meet

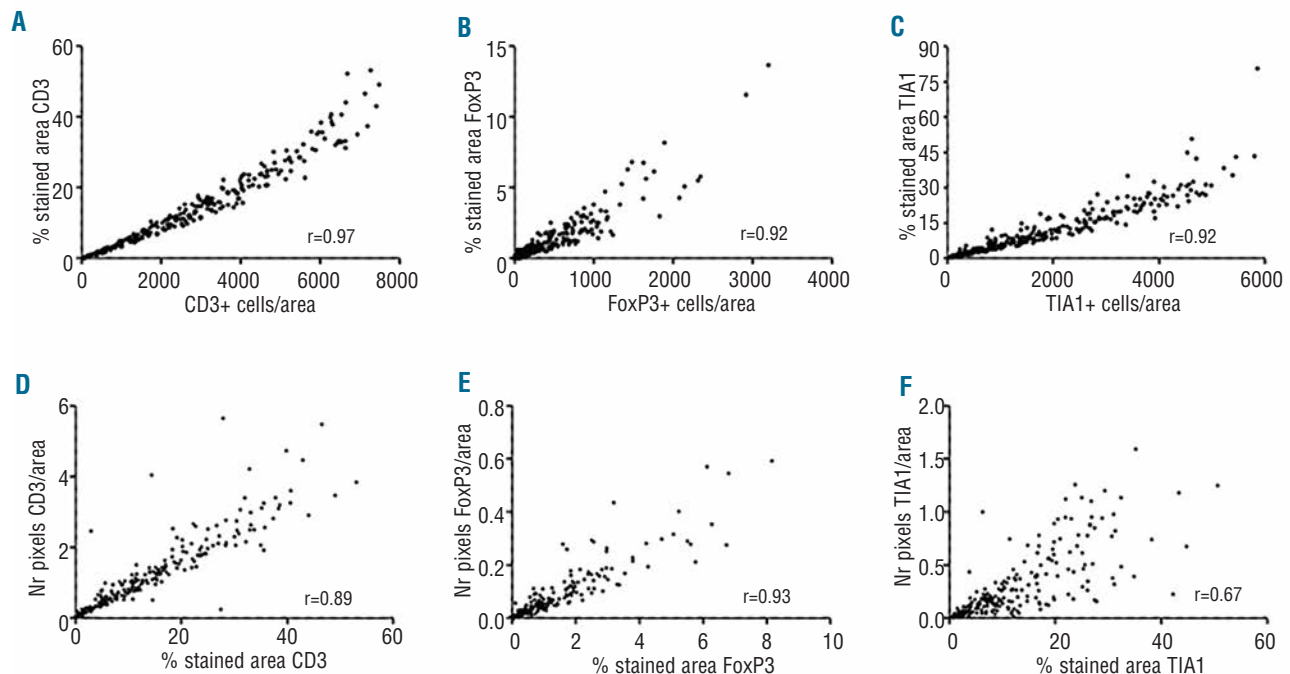


Figure 1. Density correlations for CD3, FoxP3 and TIA1. (A-C) Correlation plots between the number of positive cells per area of viable tumor and % stained area for each case as quantified by the Ariol system. (D-F) Correlation plots between the % stained area for each case as quantified by the Ariol system and the number of positive pixels per area analyzed as quantified by the Panoramic Viewer system. Pearson r values are provided.

these criteria. The most common reason for exclusion was insufficient core size or degradation.

The following clinical variables were predictive of worse overall survival by univariate analysis (*Online Supplementary Table S4*): age >60 years, stage III-IV, Eastern Cooperative Oncology Group performance status ≥ 2 and not achieving complete remission. Patients with a high International Prognostic Index had a lower probability of survival. Using the rpart package within the R software and a recursive splitting algorithm according to the criteria described previously, no cutpoint discrimination could be established for the studied biomarkers that helped to discriminate patients with different overall survivals.

With regards to progression-free survival, patients with stage III-IV disease, Eastern Cooperative Oncology Group performance status ≥ 2 , ≥ 2 extranodal areas and International Prognostic Index ≥ 3 had a lower probability of survival. Patients treated at St. Bartholomew's Hospital also had a lower progression-free survival rate compared to the ones treated at the Portuguese Institute of Oncology. Using the rpart package we defined single cut-points for CD3 and FoxP3 that segregated patients with different cell densities and progression-free survivals.

Although the number of pixels/area for CD3 has a log-rank *P*-value slightly off the established 0.01 limit for significance, the robustness of the split is confirmed by the corresponding 2x2 table in Figure 2. For both biomarkers, patients with a higher cell density had a lower progression-free survival after R-CHOP (Figures 2 and 3).

A Cox regression analysis was applied to model progression-free survival using the R-CHOP dataset. All clinical variables significantly associated with progression-free survival on univariate analysis (described above) together with the categorical data for CD3 and FoxP3 were included in the model. Using both backward and forward stepwise methods, the variables retaining independence for prediction of progression-free survival were stage III-IV disease and CD3 quantified as percentage stained area of viable tumor (*Online Supplementary Table S4*). FoxP3 did not remain as prognostically significant on multivariate analysis.

Since we used different parameters (positive cells/area of viable tumor, % stained area and number of brown pixels/area) to quantify each of the biomarkers, rpart retrieved different discriminatory cutpoints for each. Presuming that cohort discrimination is truly based on a

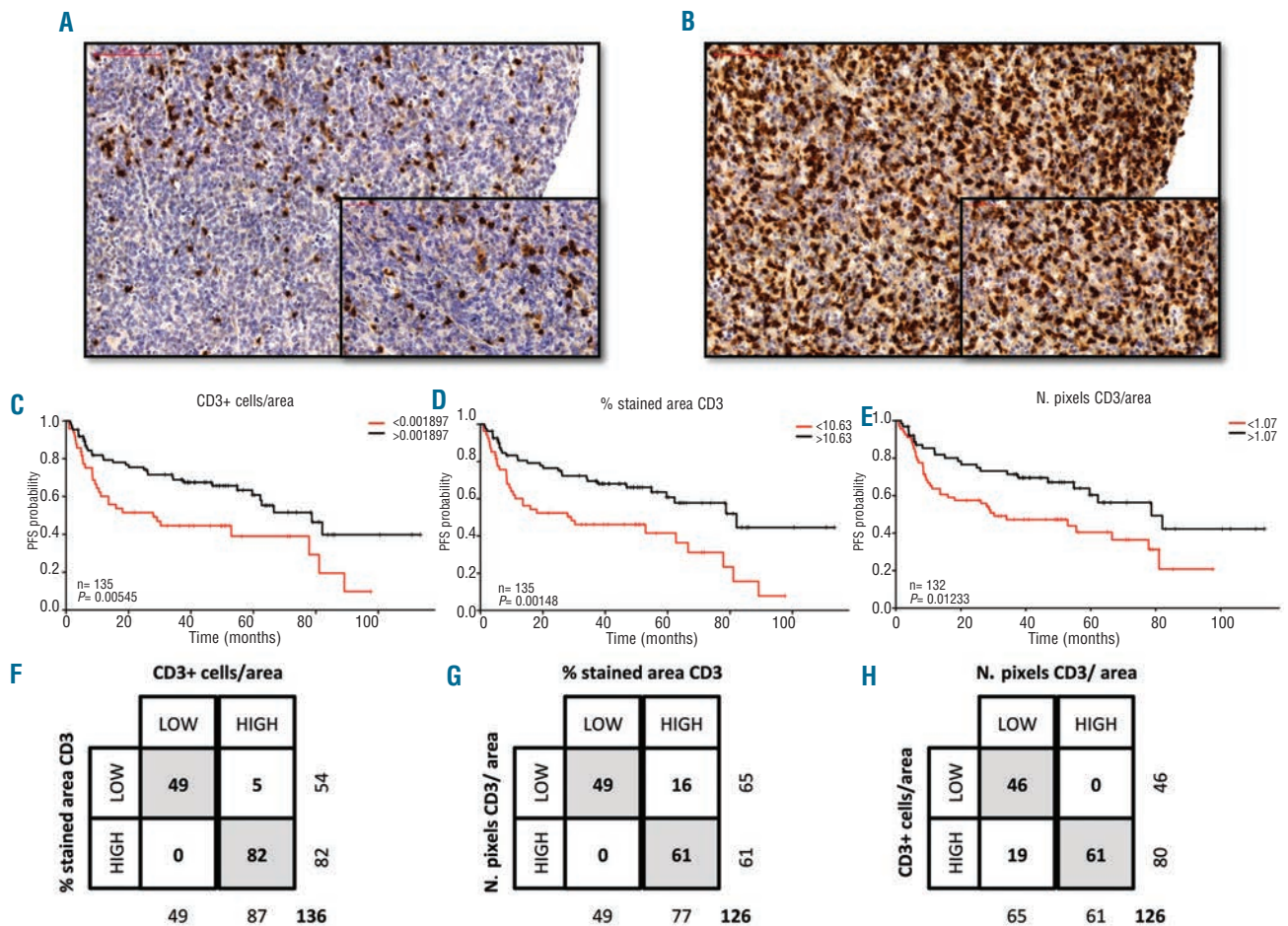


Figure 2. CD3 expression and outcome. (A-B) Representative examples of low (A) and high (B) expression of CD3 in DLBCL. Magnification x20 and x40. CD3 expression is shown by HRP-DAB immunostaining. (C-E) Kaplan Meier progression-free survival (PFS) analysis of patients based on number of low or high CD3-expressing cells/ μm^2 (C), % stained area of CD3 (D) and number of stained pixels of CD3 per area (E). Log-rank *P* values and number of cases analyzed are provided. (F-H) Contingency tables representing the distribution of patients into low and high subgroups for each of the analyses. The total number of cases represented was dependent on the availability of appropriate results for each of the analyses.

biological impact of the immune cell infiltration and not a reflection of a methodology-induced bias, we hypothesized that the agreement in allocation of each individual patient to high and low subgroups for each CD3 and FoxP3 analysis would be high. To test this hypothesis, we plotted consensus matrices and explored the degree of allocation agreement for each subgroup. As can be appreciated from Figures 2F-H and 3F-H, allocation consistency was extremely good (Fisher exact test $P < 0.0001$ for each presented 2x2 table), even when the Ariol and Panoramic Viewer results were compared. This, in our opinion, suggests that CD3 and FoxP3 are potential predictors of progression-free survival after R-CHOP treatment and validates the training-validation method as an adequate method for cutpoint selection in continuous data with unknown demographic distribution.

Discussion

Developing robust prognostic biomarkers that are able to discriminate patients treated with R-CHOP is a research

priority in DLBCL. As has been suggested by a number of gene expression studies,^{3,6} the microenvironment has the potential to provide some such biomarkers. However, a number of obstacles need to be acknowledged: (i) biomarkers need to be easily studied in any laboratory; (ii) validation processes need to be robust and performed in the context of multi-institutional projects and clinical trials; (iii) microarray technology is still reserved to the research setting and lacks well standardized procedural protocols and analysis; (iv) only a few markers selected for the gene expression profiling studies mentioned have been evaluated by independent groups using different methodologies; (v) functional studies, the only ones able to provide definitive biological explanations for the prognostic impact of specific molecules, are difficult to perform in the context of the lymphoma microenvironment.

Although IHC is highly attractive as a diagnostic and prognostic tool, the results of published IHC analyses are contradictory. This is believed to be largely related to the use of inconsistent methodologies and the fact that manual scoring of IHC staining is difficult to standardize,¹⁶ particularly when large areas of tissue are analyzed. Semi-auto-

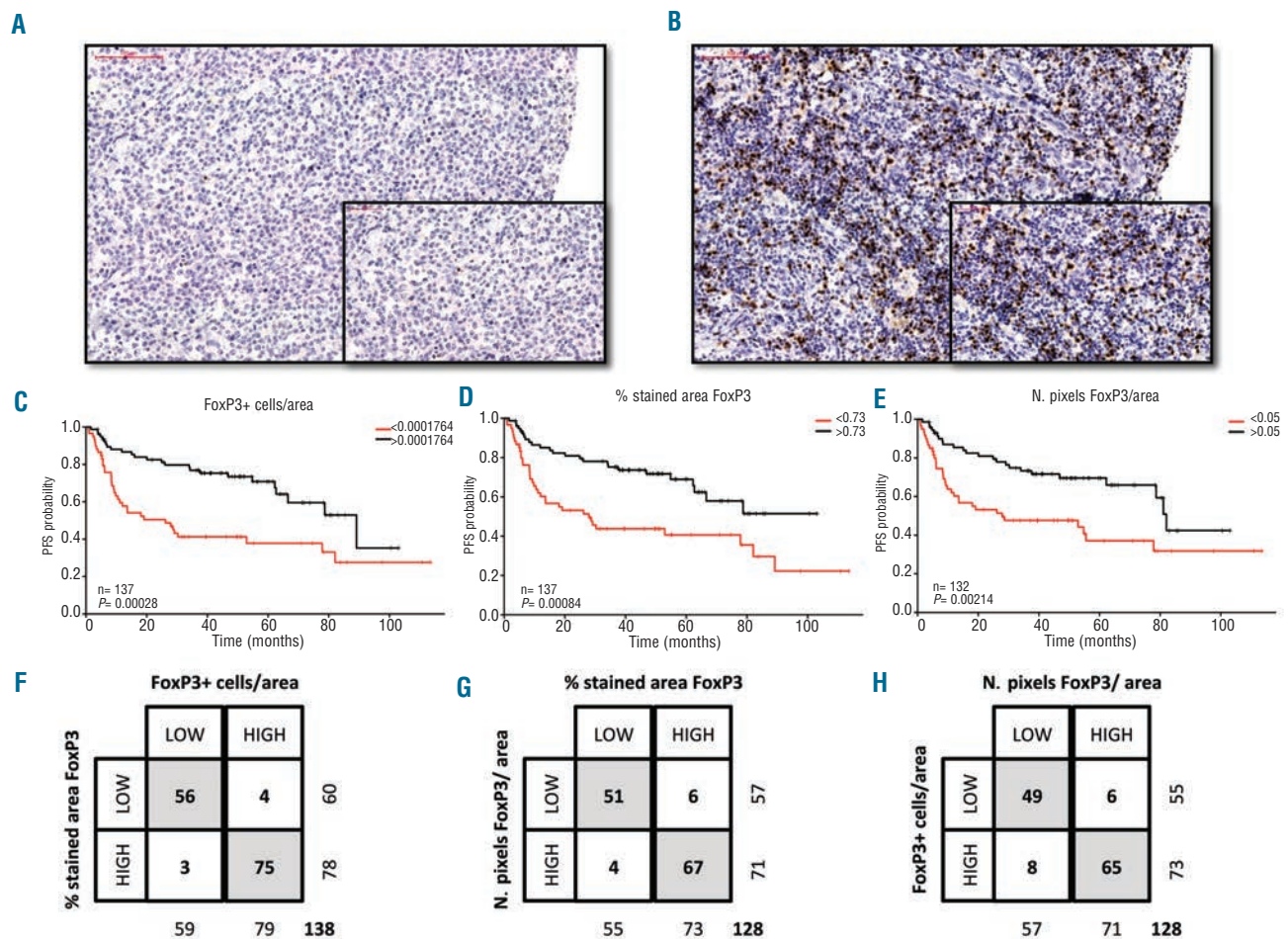


Figure 3. FoxP3 expression and outcome. (A-B) Representative examples of low (A) and high (B) expression of FoxP3 in DLBCL. Magnification x20 and x40. FoxP3 expression is shown by HRP-DAB immunostaining. (C-E) Kaplan Meier progression-free survival (PFS) analysis of patients based on number of low or high FoxP3 expressing cells/ μm^2 (C), % stained area of FoxP3 (D) and number of stained pixels of FoxP3 per area (E). Log-rank P values and number of cases analyzed are provided. (F-H) Contingency tables representing the distribution of patients into low and high subgroups for each of the analyses. The total number of cases represented was dependent on the availability of appropriate results for each of the analyses.

mated image analysis is available, is ideal for scoring vast areas more representative of the microenvironment, and could help in improving scoring reproducibility. However each system has its own features and scripts for analysis and it is unknown how robust results are when obtained from different automated systems.

This study compares two different methods of semi-automated analysis for IHC staining using the microenvironment of DLBCL as a model and examining a larger number of patients and larger area of diagnostic formalin-fixed, paraffin-embedded tissue per patient than any published work to date. Although our hypothesis was that using different systems, and hence different methodologies of analysis, would decrease reproducibility and demonstrate that such approach is unreliable, proving the opposite would support automated IHC as the way forward for assessing cell infiltration in large areas of tissue using tissue microarrays.

We, therefore, compared two image analysis systems developed by independent companies (Ariol and Panoramic Viewer), three image analysis methods (absolute numbers, percentage area and numbers of stained pixels), and as a secondary aim examined the prognostic impact of individual elements of the immune microenvironment in a cohort representative of the current clinical scenario in DLBCL. To our knowledge this approach has not been previously been undertaken for IHC studies in lymphoma.

We demonstrated that the computerized results are highly reproducible, even when comparing different variables analyzed for each biomarker, such as cell density and percentage of area stained. Correlation data for the two Ariol measurements were extremely consistent, even for cytoplasmic proteins such as TIA1 or CD68. This was a surprising result given the difficulties in training the Ariol system to enumerate single cytotoxic T cells or macrophages. Inter-observer variability for manual counting for macrophages is known to be very high. Even more surprising was the finding of highly acceptable reproducibility between the data retrieved from the Ariol and the Panoramic Viewer. Correlations were excellent for T-cell surface markers, but more modest for cytoplasmic proteins.

Our data are in agreement with those of a recent validation study promoted by the Lunenburg Lymphoma Biomarker Consortium (LLBC).¹⁵ This study, conducted by highly experienced hematopathologists, found only low to moderate agreement in manual scoring of T cells and macrophage markers when four to five scoring categories were used. However, comparison of automated analyses set up by two different operators in two different instruments using the same methodology was highly reproducible for T-cell markers. This is an important finding suggesting that operator-induced bias is not as relevant as had been thought previously and should not prevent researchers from using this methodology. The current study adds to this by analyzing a large population of 309 patients and by comparing the computerized quantifications for a larger panel of markers, particularly including cytoplasmic proteins such as CD68. Moreover, we performed comparisons across the results retrieved from two semi-automated systems developed by different companies and three different assays for each stain.

Our own and the LLBC data indicate that semi-automated systems of IHC analysis add the required robustness to IHC prognostic studies in an operator-independent manner

and should be used in the future instead of manual analysis. Comparing manual and automated results was not the scope of this project and studies such as those of the LLBC highlight that manual scoring is highly variable and hence inadequate for outcome prediction studies. This is being explored further in follicular lymphoma by the LLBC.

Finally we explored the outcome potential of microenvironment biomarkers as assessed using the semi-automated systems in a representative dataset of 161 uniformly R-CHOP-treated DLBCL patients. Whereas the clinical variables included in the International Prognostic Index and achieving a complete remission after R-CHOP were predictive of overall survival, none of the biomarkers studied was, potentially reflecting the impact of salvage therapy. However we were able to validate CD3 and FoxP3 as predictors of progression-free survival. Patients with higher densities of CD3 and FoxP3 had lower risks of relapse after R-CHOP treatment. Although studies highlighting CD3 as a potential marker in lymphoma are limited,¹⁷ previous IHC analyses predominantly showed that a high infiltration of FoxP3 improves patients' survival.¹⁸⁻²⁴ Although functional studies pointed towards a negative prognostic impact for regulatory T cells in lymphoma,²⁵⁻²⁷ it is possible that regulatory T cells directly suppress malignant B cells or counteract tumor-supporting T cells. Notably, we saw no evidence that CD68 expression, determined by any method, was prognostic in our R-CHOP-treated DLBCL patients. We now face a challenge of which methodology to propose for future studies. It could be argued that discriminating prognostic cohorts based on analysis of immunohistochemical markers reflects only bias arising from the methodology. However, our use of multiple methodologies, technologies and markers, all suggesting a positive impact of increased immune cell infiltrate, supports the hypothesis that it is representative of a real biological effect: that increased immune infiltrate leads to improved outcome. Moreover, it is another finding indicating that computerized analysis is a robust method for IHC analysis.

Clearly, the results reported here are exploratory. Now that we have established that semi-automated systems are the tool of choice for analysis of IHC biomarkers, larger validation studies are required. First, it is imperative to conduct an intergroup analysis of all automated systems available, devise an agreed methodological approach and select laboratories that would be responsible for similar analyses under clinical trials in which other established molecular and cytogenetic prognostic factors are being investigated. This, in our opinion, would definitively resolve the issue of whether the lymphoma microenvironment plays a role in outcome prediction in DLBCL and other lymphomas and bring forward methods to incorporate such biomarkers into clinical practice.

Acknowledgments

The authors would like to thank the Portuguese Foundation for Science and Technology for providing grant support to RC with a Doctoral fellowship (SFRH/BD/68462/2010) and to the National Cancer Institute (P01 CA81538) and Cancer Research UK (C1574/A6806) for program grants to JGG.

Authorship and Disclosures

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. Coiffier B, Lepage E, Briere J, et al. CHOP chemotherapy plus rituximab compared with CHOP alone in elderly patients with diffuse large-B-cell lymphoma. *N Engl J Med.* 2002;346(4):235-242.
2. Coiffier B, Thieblemont C, Van Den Neste E, et al. Long-term outcome of patients in the LNH-98.5 trial, the first randomized study comparing rituximab-CHOP to standard CHOP chemotherapy in DLBCL patients: a study by the Groupe d'Etudes des Lymphomes de l'Adulte. *Blood.* 2010;116(12):2040-2045.
3. Alizadeh AA, Eisen MB, Davis RE, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature.* 2000;403(6769):503-511.
4. Rosenwald A, Wright G, Chan WC, et al. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N Engl J Med.* 2002;346(25):1937-1947.
5. Monti S, Savage KJ, Kutok JL, et al. Molecular profiling of diffuse large B-cell lymphoma identifies robust subtypes including one characterized by host inflammatory response. *Blood.* 2005;105(5):1851-1861.
6. Lenz G, Wright G, Dave SS, et al. Stromal gene signatures in large-B-cell lymphomas. *N Engl J Med.* 2008;359(22):2313-2323.
7. Hasselblom S, Hansson U, Sigurdardottir M, Nilsson-Ehle H, Ridell B, Andersson PO. Expression of CD68+ tumor-associated macrophages in patients with diffuse large B-cell lymphoma and its relation to prognosis. *Pathol Int.* 2008;58(8):529-532.
8. Hasselblom S, Sigurdardottir M, Hansson U, Nilsson-Ehle H, Ridell B, Andersson PO. The number of tumour-infiltrating TIA-1+ cytotoxic T cells but not FOXP3+ regulatory T cells predicts outcome in diffuse large B-cell lymphoma. *Br J Haematol.* 2007;137(4):364-373.
9. Tzankov A, Meier C, Hirschmann P, Went P, Pileri SA, Dimhofer S. Correlation of high numbers of intratumoral FOXP3+ regulatory T cells with improved survival in germinal center-like diffuse large B-cell lymphoma, follicular lymphoma and classical Hodgkin's lymphoma. *Haematologica.* 2008;93(2):193-200.
10. Muris JJ, Meijer CJ, Cillessen SA, et al. Prognostic significance of activated cytotoxic T-lymphocytes in primary nodal diffuse large B-cell lymphomas. *Leukemia.* 2004;18(3):589-596.
11. Meyer PN, Fu K, Greiner T, et al. The stromal cell marker SPARC predicts for survival in patients with diffuse large B-cell lymphoma treated with rituximab. *Am J Clin Pathol.* 2011;135(1):54-61.
12. Wada N, Zaki MA, Hori Y, et al. Tumour-associated macrophages in diffuse large B-cell lymphoma: a study of the Osaka Lymphoma Study Group. *Histopathology.* 2012;60(2):313-319.
13. Marmey B, Boix C, Barbaroux JB, et al. CD14 and CD169 expression in human lymph nodes and spleen: specific expansion of CD14+CD169- monocyte-derived cells in diffuse large B-cell lymphomas. *Hum Pathol.* 2006;37(1):68-77.
14. Ihaka R, Gentleman R. R: a language for data analysis and graphics. *J Comput Graph Stat.* 1996;5(3):299-314.
15. Sander B, de Jong D, Rosenwald A, et al. The reliability of immunohistochemical analysis of the tumor microenvironment in follicular lymphoma: a validation study from the Lunenburg Lymphoma Biomarker Consortium. *Haematologica.* 2014;99(4):715-725.
16. de Jong D, Rosenwald A, Chhanabhai M, et al. Immunohistochemical prognostic markers in diffuse large B-cell lymphoma: validation of tissue microarray as a prerequisite for broad clinical applications--a study from the Lunenburg Lymphoma Biomarker Consortium. *J Clin Oncol.* 2007;25(7):805-812.
17. Wahlin BE, Sundstrom C, Holte H, et al. T cells in tumors and blood predict outcome in follicular lymphoma treated with rituximab. *Clin Cancer Res.* 2011;17(12):4136-4144.
18. Farinha P, Al-Tourah A, Gill K, Klasa R, Connors JM, Gascoyne RD. The architectural pattern of FOXP3-positive T cells in follicular lymphoma is an independent predictor of survival and histologic transformation. *Blood.* 2010;115(2):289-295.
19. Lee AM, Clear AJ, Calaminici M, et al. Number of CD4+ cells and location of forkhead box protein P3-positive cells in diagnostic follicular lymphoma tissue microarrays correlates with outcome. *J Clin Oncol.* 2006;24(31):5052-5059.
20. Alvaro T, Lejeune M, Salvado MT, et al. Outcome in Hodgkin's lymphoma can be predicted from the presence of accompanying cytotoxic and regulatory T cells. *Clin Cancer Res.* 2005;11(4):1467-1473.
21. Alvaro T, Lejeune M, Salvado MT, et al. Immunohistochemical patterns of reactive microenvironment are associated with clinicobiologic behavior in follicular lymphoma patients. *J Clin Oncol.* 2006;24(34):5350-5357.
22. Sweetenham JW, Goldman B, LeBlanc ML, et al. Prognostic value of regulatory T cells, lymphoma-associated macrophages, and MUM-1 expression in follicular lymphoma treated before and after the introduction of monoclonal antibody therapy: a Southwest Oncology Group Study. *Ann Oncol.* 2010;21(6):1196-1202.
23. Wahlin BE, Aggarwal M, Montes-Moreno S, et al. A unifying microenvironment model in follicular lymphoma: outcome is predicted by programmed death-1--positive, regulatory, cytotoxic, and helper T cells and macrophages. *Clin Cancer Res.* 2010;16(2):637-650.
24. Greaves P, Clear A, Coutinho R, et al. Expression of FOXP3, CD68, and CD20 at diagnosis in the microenvironment of classical Hodgkin lymphoma is predictive of outcome. *J Clin Oncol.* 2013;31(2):256-262.
25. Yang ZZ, Novak AJ, Stenson MJ, Witzig TE, Ansell SM. Intratumoral CD4+CD25+ regulatory T-cell-mediated suppression of infiltrating CD4+ T cells in B-cell non-Hodgkin lymphoma. *Blood.* 2006;107(9):3639-3646.
26. Yang ZZ, Novak AJ, Ziesmer SC, Witzig TE, Ansell SM. Attenuation of CD8(+) T-cell function by CD4(+)CD25(+) regulatory T cells in B-cell non-Hodgkin's lymphoma. *Cancer Res.* 2006;66(20):10145-10152.
27. Galani IE, Wendel M, Stojanovic A, et al. Regulatory T cells control macrophage accumulation and activation in lymphoma. *Int J Cancer.* 2010;127(5):1131-1140.