Introduction

Follicular lymphoma (FL) is a B cell lymphoproliferative disorder which usually has an indolent course with median overall survival (OS) of almost 20 years, especially after the introduction of rituximab (1-5). The current standard of care management is immunochemotherapy for those patients who require treatment, with both rituximab and obinutuzumab-based therapy leading to prolonged progression-free survival (PFS). Long term follow-up of the FOLL05 trial has shown an 8-year PFS of 48% without rituximab maintenance, and patients treated with rituximab maintenance in the PRIMA trial had a median PFS of 10.5 years (6,7). Despite these encouraging long-term results, a subset of patients experiences early progression after standard therapy. Recent evidence has demonstrated that 10–20% of patients who are treated with immunochemotherapy experience relapse within 24 months of diagnosis and these patients have a significantly worse outcome compared to those who do not. While well-established clinical risk factors allow to predict outcomes, risk-stratification is improved when considering molecular attributes such as gene mutations or gene expression profiles. Similarly, functional imaging allows to estimate the burden of disease prior to initiating therapy and provides useful information at the end of frontline therapy, allowing to distinguish those patients who achieve a complete metabolic response and have superior survival from patients who have residual, metabolically active lymphoma. Ongoing work focuses on predicting early progression with greater accuracy, which would enable the study of risk-adapted management strategies. Increasingly, risk indices are recognized to have differential prognostic impact based on the specific treatment under consideration. It is conceivable that future risk stratification will not only provide an estimate of prognosis, but also inform the most appropriate treatment choice. Herein, we describe the existing and emerging risk-stratification tools, based on clinical, molecular and imaging attributes, and review their ability to predict early progression.

Keywords: Follicular lymphoma; outcome prediction; risk stratification; progression.
Early progression

In recent years, early progression has been recognized as a critical determinant of patient outcomes. It was initially recognized that patients who experienced progression during initial immunochemotherapy or maintenance phase (considered immunochemotherapy resistant) had a poorer outcome compared to those who responded to treatment (8). In a cohort of 132 patients, 22 (16.7%) were deemed to have lymphoma that was resistant to immunochemotherapy, and 8 of these 22 patients (36%) were subsequently found to have transformed disease. The PFS and OS were not reached in the immunochemotherapy responsive group (median follow-up 33 months) and were 17 months and 47 months, respectively, in those who had resistant FL, demonstrating a poorer outcome in these patients who progressed during initial immunochemotherapy (8).

Casulo et al. examined patients in the National LymphoCare Study (NLCS) who received treatment for FL, focusing on patients treated with R-CHOP (rituximab, cyclophosphamide, doxorubicin, vincristine and prednisone) (9). In this cohort, 110 patients out of 588 (19%) experienced early progression within 24 months of diagnosis (POD24). The 2- and 5-year OS in the POD24 cohort were 68% and 50%, respectively, compared to 97% and 90% in the reference group, with a hazard ratio (HR) of 7.2. The increased risk of the POD24 cohort was maintained after adjustment for the FLIPI score. POD24 was also significantly associated with poor survival in the comparison with patients who had progression of disease after 24 months. In a validation cohort from the University of Iowa and Mayo Clinic Molecular Epidemiology Resource, also treated with R-CHOP, 26% of patients experienced POD24, with 2- and 5-year OS being 64% and 34% compared to 98% and 94%, respectively, in the reference group. An exploratory analysis of the R-CVP (rituximab, cyclophosphamide, vincristine, prednisone) and R-Flu (rituximab, fludarabine) groups in the NLCS cohort demonstrated similarly an association between POD24 and decreased survival.

Patients treated with bendamustine and rituximab who experienced POD24 had histologically proven transformed disease at progression, suggesting that bendamustine may reduce the incidence of progression but not the occurrence of transformation to aggressive histology. A retrospective analysis of the GALLIUM study examined 1,202 patients who were randomized to obinutuzumab or rituximab treatment combined with either bendamustine, CVP or CHOP, and found that POD24, defined as progressive disease within 24 months of randomization, occurred in 13% of patients (11). Fifty-seven out of 601 patients treated with obinutuzumab experienced POD24 (10%), compared to 98 out of 601 (16%) patients treated with rituximab-based therapy. The risk reduction of POD24 events was 46% with obinutuzumab, relative to rituximab. The risk of mortality associated with early progression was influenced by the timing of progression, with the highest risk seen in those patients with very early progression (within 6 months). The finding that early progression events—with and without evidence of histological transformation—confer poor survival if they occur during or after immunochemotherapy, has been validated in several additional studies (12-15).

This conclusion needs, however, to be somewhat nuanced in the light of the various treatment approaches that can be considered for FL. A study analysing early progression (within 12 months) in patients from the University of Iowa and Mayo Clinic Molecular Epidemiology Resource found that patients with early progression who were initially treated with observation or rituximab monotherapy had significantly less poor outcome compared to those who were treated with immunochemotherapy (14). Hence, the intensity of the treatment received prior to progression appears to be correlated with the risk of mortality after early progression. An analysis of Swiss Group for Clinical Cancer Research (SAKK) trials focused on chemotherapy-free regimens in patients who were previously untreated (16). This analysis included 333 patients, of whom 256 received rituximab alone and 77 patients who received rituximab with lenalidomide for six months. Twenty-seven percent of 318 evaluable patients experienced POD24 with 5-year and 10-year OS of 69% and 59% compared to 97% and 77% in patients without POD24. In patients who received rituximab only, the 5-year OS was 71% versus 93% in patients with POD24 and without, respectively. A further nuance relates to the detection of early progression, as Bitansky et al. reported that POD24 did not predict OS in patients in whom progression was documented incidentally on imaging (17). Nonetheless, especially for patients who are...
Clinical risk tools

**Follicular Lymphoma International Prognostic Index (FLIPI)**

The FLIPI was initially developed because the International Prognostic Index (IPI) was found not to be as sensitive for low grade lymphoma, given that it was created with high grade lymphomas in mind. Retrospective data were analyzed to identify characteristics associated with poor outcomes in FL patients, narrowed down to five variables which would be user-friendly in the clinical setting (18). These included age ≥60, Ann-Arbor stage III–IV, hemoglobin level <120 g/L, lactate dehydrogenase (LDH) > upper limit of normal and more than 4 nodal sites. Based on these five variables, three risk categories were formed: low [0–1], intermediate [2–3], and high [4–5]. The distribution of low, intermediate and high-risk was 36%, 37% and 27%, respectively.

As the FLIPI was developed based on a patient cohort diagnosed between 1985 and 1992, the treatment was heterogeneous between different centres and varied over time. However, there have been subsequent analyses which have validated the FLIPI in context of treatment with rituximab-containing treatment. In the German Low Grade Lymphoma Study Group (GLSG), the 2-year time to treatment failure was lower in the high-risk group (67%) compared to low-risk (92%) and intermediate-risk (90%) patients (19). In a prospective study that compared R-CVP to CVP (20), a multivariate analysis showed that the FLIPI and assignment to the experimental arm were the only variables that were significantly associated with time to progression. In a prospective, observational cohort study by the NLCS group, which included 2,192 assessable patients accrued between 2004 and 2007, the percentages of patients assigned to the low, intermediate and high-risk categories were 35%, 30% and 35%, respectively (21). Sixty-eight percent were treated with rituximab-based therapy, whereas 17% were observed and 15% were treated with non-rituximab based therapy. Intermediate and high-risk FLIPI were associated with HRs for OS of 2.3 and 7.1, respectively. More limited data are available that correlate risk stratification using the FLIPI specifically with early progression. Jurinovic et al. reported that the FLIPI had a sensitivity of 70–78% but a specificity of only 56–58% to predict POD24 in patients treated with R-CHOP or R-CVP (15), suggesting that many patients assigned into the high-risk FLIPI category may not ultimately experience POD24.

**Follicular Lymphoma International Prognostic Index 2 (FLIPI2)**

Subsequently, a second prognostic model was put forth, called FLIPI2, which was developed with a prospective cohort of newly diagnosed FL patients in the post rituximab era and using PFS rather than OS as an endpoint (22). This excluded patients whose initial approach was watch and wait as the endpoint of PFS was deemed less relevant to them. The final analysis included 832 patients. Five variables were selected to be included in the final model: beta-2 microglobulin (B2M) above the upper limit of normal, longest dimension of single lymph node longer than 6 cm, bone marrow involvement, hemoglobin lower than 120 g/L and age greater than 60. Patients with 0 risk factors were considered low-risk, 1–2 risk factors considered intermediate and 3–5 risk factors considered high-risk. The FLIPI2 was subsequently validated by other studies (23-25). Overall, the FLIPI2 appears to have similar discriminatory ability to predict PFS, when compared to the FLIPI (22,26).

**PRIMA-Prognostic Index (PRIMA-PI)**

A more recent prognostic index that was developed and trained on the PRIMA study patients is referred to as the PRIMA-PI (26). Briefly, the PRIMA study included FL patients requiring treatment, treated with immunochemotherapy, with or without rituximab maintenance for 2 years. The PRIMA-PI had only two variables: B2M and bone marrow involvement. Those with B2M ≥3 mg/L were considered high-risk. Patients with B2M ≤3 mg/L and with bone marrow involvement were found to have intermediate-risk, and patients with B2M ≤3 mg/L and with no bone marrow involvement had low-risk. The 5-year PFS rates were 37%, 55% and 69% in the high, intermediate and low-risk categories, respectively. The study further calculated FLIPI and FLIPI2 scores on the training cohorts and compared the three prognostic models. The PRIMA-PI more evenly divided patients into three categories whereas there were very few patients who were low-risk using the FLIPI2 (6%). In the PRIMA data, the prognostic performance of the PRIMA-PI appeared improved, compared to the FLIPI and the FLIPI2, but the
PRIMA-PI and the FLIPI performed similarly in a pooled validation cohort. The percentages of patients with an event within 24 months of diagnosis were 14%, 21% and 38% in the low, intermediate and high-risk PRIMA-PI groups, respectively.

In a subsequent analysis based on the RELEVANCE trial, the PRIMA-PI was found to divide patients into equal proportions of low, intermediate or high-risk patients, whereas the FLIPI and the FLIPI2 classified only 15% and 8% of patients into the low-risk group, which is suboptimal given that many FL patients experience favorable outcomes (27). The PRIMA-PI and the FLIPI performed similarly with regards to separating patient groups based on PFS. Interestingly, none of the indices was significantly associated with PFS in the group of patients treated with rituximab and lenalidomide, suggesting that distinct prognostic tools should be considered in these patients. On the other hand, the PRIMA-PI allowed the identification of high-risk patients in Nordic Lymphoma Group trials where patients received rituximab, with or without interferon (28).

**Follicular Lymphoma Evaluation Index (FLEX)**

The FLEX score was developed to improve the identification of high-risk patients (29). This predictive model was trained in 1,202 patients from the GALLIUM trial and validated in patient data from the SABRINA trial. The FLEX score used nine clinical variables to separate patients into low-risk (score 0–2, 64% of patients in training cohort) or high-risk groups (score 3–9, 36% of patients), with improved distinction between low and high-risk categories when compared to the FLIPI, the FLIPI2 or the PRIMA-PI. The nine variables consisted of: male sex, high sum of the products of lesion dimensions, grade 3A, more than 2 extranodal sites, ECOG performance status greater than 1, hemoglobin less than 12 g/dL, elevated B2M, natural killer cell count in peripheral blood lower than 100/μL and elevated serum LDH. In the training cohort, the sensitivity to predict POD24 was highest for the PRIMA-PI (69%), followed by the FLEX score (60%) and the FLIPI/FLIPI2 (53%). The specificity on the other hand was highest for the FLEX score (68%), followed by the FLIPI/FLIPI2 (59%) and the PRIMA-PI (47%).

An overview of the proportions of patients experiencing POD24 within clinical risk categories is shown in Figure 1, and an overview of prognostic models that identify patients at increased risk of early progression is shown in Table 1. While risk scores continue to be refined, neither of the available indices has thus far had a definitive role in altering clinical management, mostly because their accuracy to identify high-risk situations remains imperfect. They also do not capture biological dimensions that can be important determinants of patient outcomes.

**Individual genetic alterations and genomic properties underlying early progression**

The genomic landscape of FL is characterized by highly recurrent mutations in genes encoding proteins involved in several key pathways, with the most frequent mutations affecting epigenetic modifiers (30-34). Yet, unlike other indolent B-cell malignancies such as chronic lymphocytic leukemia (CLL), in which ascertainment of specific gene mutations and/or deletions informs clinical decision making, there is a paucity of data that reliably associates individual gene mutations seen in FL with patient outcomes. Nonetheless, such mutations may identify distinct clinical trajectories.

For example, TP53 gene alterations are reported to be present in ~5% of FL patients at the time of diagnosis and have generally been associated with inferior outcome with early progression and transformation of disease in the pre- and post-rituximab eras (12,35-38). Interestingly, subclonal TP53 mutations can be detected in a significant proportion of FL cases, which led to an overall percentage of mutated samples of 23% in the SWOG S0016 trial, in which patients were treated with R-CHOP or CHOP plus 131-iodine tositumomab (35). Such subclonal events appeared to have similar prognostic information when compared to clonally dominant mutations, although this observation deserves further validation.

FL samples harbor frequent gain-of-function mutations of EZH2, encoding a histone methyltransferase. In the PRIMA study, EZH2 mutations or gains were found in 37% of samples analyzed, which was demonstrated to be associated with a favorable PFS. In terms of early progression, 31% and 15% of patients had progression within 2 years in the non-altered and altered groups, respectively (39). Similar findings were also seen in a study from the Lunenburg Lymphoma Biomarker Consortium (40). In the GALLIUM trial, 22% of samples were found to be EZH2-mutated. Those without EZH2 mutation who were treated with CHOP/CVP were found to have an unfavorable prognosis, but longer PFS when treated with bendamustine (36). Beyond being a prognostic biomarker, EZH2 mutation status is emerging...
Figure 1 Number of follicular lymphoma patients experiencing POD24 in risk categories defined by the FLIPI, the m7-FLIPI, the POD24-PI, the PRIMA-PI or the FLEX scores. As the reported studies had varying sample sizes, the numbers shown here were normalized to represent numbers of patients out of 100. For each study, only the results from the training cohorts are shown. In Jurinovic et al., POD24 was defined as progression or relapse within 24 months of first-line treatment (15). Bachy et al. reported EFS24, defined as event-free survival within 24 months of diagnosis (26). Mir et al. defined POD24 as progression or disease-related death within 24 months of randomization (29).

Table 1 Follicular lymphoma risk assessment tools and their association with POD24

<table>
<thead>
<tr>
<th>Components</th>
<th>Risk groups</th>
<th>Survival†</th>
<th>POD24</th>
</tr>
</thead>
<tbody>
<tr>
<td>FLIPI (15,18,29)</td>
<td>Age; Ann-Arbor stage; Hb; LDH; nodal sites</td>
<td>L: 0–1 points; I: 2–3 points; H: 4–5 points</td>
<td>L: 92%; I: 90%; H: 67% (2-year TTF)</td>
</tr>
<tr>
<td>FLIPI2 (22)</td>
<td>B2M; longest dimension of single lymph node; BMI; age</td>
<td>L: 0 points; I: 1–2 points; H: 3–5 points</td>
<td>L: 91%; I: 69%; H: 51% (3-year PFS)</td>
</tr>
<tr>
<td>FLEX (29)</td>
<td>Male sex; sum of lesion dimension; grade 3A; extra nodal sites; ECOG; Hb; B2M; NK cell count; LDH</td>
<td>L: 0–2 points; H: 3–9 points</td>
<td>L: 86%; H: 68% (3-year FFS)</td>
</tr>
<tr>
<td>m7-FLIPI (15,36)</td>
<td>FLIPI; ECOG; mutation status of 7 genes (ARID1A, CARD11, CREBBP, EP300, EZH2, FOXO1, MEF2B)</td>
<td>L: score &lt;0.8; H: score &gt;0.8</td>
<td>L: 77%; H: 38% (5-year FFS)</td>
</tr>
<tr>
<td>POD24-PI (15)</td>
<td>High risk FLIPI; mutation status of 3 genes (EP300, EZH2, FOXO1)</td>
<td>L: score &lt;0.71; H: score &gt;0.71</td>
<td>L: 77%; H: 50% (5-year FFS)</td>
</tr>
</tbody>
</table>

†, survival data from initial paper describing the prognostic model. L, low risk; I, intermediate risk; H, high risk; BMI, bone marrow involvement; B2M, beta-2 microglobulin; Hb, hemoglobin; LDH, lactate dehydrogenase; PFS, progression free survival; FFS, failure free survival; TTF, time to treatment failure; NK, natural killer; ECOG, Eastern Cooperative Performance Status.

as a critical factor that predicts superior response rates to EZH2 inhibition (41). CREBBP is another gene that encodes an epigenetic modifier, and is mutated in 60–70% of FL cases (36). The type of CREBBP mutation appears to have prognostic information, with mutations in the lysine acetyltransferase domain associated with worse outcome than nonsense/frameshift mutations (42). It can be speculated that mutation ascertainment will become increasingly relevant for clinical decision-making in the future.

Beyond individual gene alterations, several studies have
described genomic properties specifically in the light of early progression. In a correlative analysis of the SWOG S0016 trial, an increased number of genetic alterations (either copy number alterations or copy-neutral loss of heterozygosity) was associated with a greater propensity of progression within 2 years (38). Similarly, the total mutation burden in coding regions of the genome has also been associated with early progression (12,43). Importantly, a study using whole-genome sequencing of paired lymphoma samples found that the clonal evolution underlying early progression was limited, at least when compared to clonal changes underlying transformation (12). This observation may suggest that the determinants of early progression are potentially present in the diagnostic tissue specimen, whereas transformation to aggressive histology may be much more difficult to predict (44).

**Clinico-genetic risk models**

Given that both clinical variables and biological annotation are associated with patient outcomes, efforts have been undertaken to combine these in so-called “clinico-genetic risk models”. The m7-FLIPI is such a model that was developed after assessing the mutation status of 74 genes in two cohorts of patients. The training cohort was derived from the German GLSG2000 trial and included 151 patients treated with R-CHOP, whereas the validation cohort was composed of samples from 107 patients from BC Cancer and treated with R-CVP (36). Seven genes were selected in an unbiased fashion using penalized Cox regression, with the strongest positive coefficients (correlating with adverse prognosis) seen with mutations of EP300 and FOXO1, and the strongest negative coefficients (favorable prognosis) observed with ARID1A and EZH2. The m7-FLIPI score was calculated by integrating the mutation of these seven genes with two clinical risk variables, namely the FLIPI and ECOG performance status, and dichotomized based on an optimal cutoff into high- and low-risk categories. While dichotomization of a continuous variable results in some information being lost, it is nonetheless a useful approach to conceptualize the clinical and biological significance of various risk categories. In the training group, the high-risk group had 5-year freedom-from-survival (FFS) of 38% versus 77% in the low-risk group. In the validation group, the 5-year FFS was 25% in the high-risk versus 68% in the low-risk group. The m7-FLIPI improved risk stratification mainly by reclassifying patients who were falsely classified as high-risk by the FLIPI into the low-risk m7-FLIPI group. The m7-FLIPI was also significantly associated with 5-year OS in both cohorts.

A recent correlative analysis of the GALLIUM trial further evaluated the prognostic utility of m7-FLIPI (45). Of the 418 patients available for analysis, 104 patients were classified as high-risk m7-FLIPI, which was associated with shorter PFS. However, the prognostic utility was largely dependent on the specific treatments that patients had received. Indeed, whereas the m7-FLIPI predicted outcome in patients treated with rituximab-based regimens, it was not prognostic for patients treated with obinutuzumab. Furthermore, it was prognostic in patients who were treated with CHOP/CVP, but not in patients receiving bendamustine. The varying outcome information contained in the m7-FLIPI appeared to be driven by EZH2 mutation status. Taken together, these data suggest the m7-FLIPI and/or EZH2 mutation status may allow to select the preferred chemotherapeutic backbone to use in combination with an anti-CD20 targeting antibody, hence be predictive, rather than prognostic biomarkers.

Given that the m7-FLIPI was not developed with the specific endpoint of POD24 in mind, a study by Jurinovic et al. examined its utility to predict early progression (15). POD24 occurred in 15% and 18% of the previously mentioned GLSG and BC Cancer cohorts, respectively, and 70–78% of POD24 cases were classified as high-risk by the FLIPI. On the other hand, 42–44% of patients without POD24 were assigned into the high-risk FLIPI category, but their FFS was comparable to patients without POD24 and low-risk FLIPI. Compared to the FLIPI, the m7-FLIPI had greater specificity to predict POD24. High-risk m7-FLIPI patients had shorter FFS even in the absence of POD24. Using the same patient cohorts and gene mutation data, a POD24-specific prognostic model was trained and included one clinical variable (the FLIPI) and the mutation status of three genes (EP300, FOXO1 and EZH2). Compared to the m7-FLIPI, the new model (POD24 Prognostic Index, or POD24-PI) had greater sensitivity to predict POD24, but lower specificity. Ultimately, the m7-FLIPI was found to have the highest accuracy and positive predictive value for POD24 and identified a smaller group of high-risk patients, when compared to the FLIPI or the POD24-PI, suggesting that it may be a suitable biomarker for further evaluation.

**Gene expression-based prediction models**

Gene expression profiling of FL tissue samples has been...
a promising technique to determine the expression levels of thousands of genes at once, and has been applied early on to FL samples. The seminal study by Dave et al. aimed to predict survival in a group of 191 patients treated with a variety of treatment approaches and defined two immune response signatures that were associated with length of survival (46). Given the emerging signal that outcome correlations are tightly related to the specific treatments received, it is not surprising that these and other immune signatures have not been universally validated. Other studies have applied gene expression profiling to dissect patient cohorts into groups of patients separated by diverging outcomes. For example, examining gene expression networks, Gentles et al. found that genes highly expressed in embryonic stem cells were associated with a higher risk of transformation and lower survival rates (47). Similarly, Brodtkorb et al. described that the expression of NF-κB pathway genes correlated with transformation in a cohort of 44 well-annotated patient samples (48). Steen et al. later extended these findings to define a high-risk group of patients identified by both high-risk FLIPI and a high-risk BTK score (49).

To define robust signatures predicting adverse outcomes, Huet et al. identified gene expression changes that were correlated with PFS in the PRIMA trial (50). They defined a gene expression model based on the expression of 23 genes that included genes involved with pathways of DNA repair, cell development, migration and immune response and that was adapted into a digital gene expression assay. In a multivariate Cox model which adjusted for rituximab maintenance and the FLIPI, those identified as high-risk by the 23 gene predictor had a 5-year PFS of 26% versus 73% in those with low-risk. Remarkably, this model appeared very robust as it could be confirmed in three independent validation cohorts. Furthermore, using an adapted gene expression predictor, Silva et al. validated the prognostic relevance of a modified “23” gene predictor in 137 patients treated with R-CVP, with or without maintenance (51). A significant association was found between high-risk assignment by the gene expression predictor and immunohistochemical expression of the FOXP1 transcription factor that had previously been shown to predict inferior outcome (52). The gene expression signature appeared to further separate outcome groups within m7-FLIPI high and low-risk categories, suggesting added benefit from assessing multiple biomarkers. A recent study by Bolen et al. demonstrated that the prognostic information obtained from the 23 gene predictor varied depending on the chemotherapy arm used, in keeping with similar observations made for the m7-FLIPI (53). The authors determined the 23 gene predictor class in 274 patients from the GALLIUM trial using RNA sequencing. They identified a chemotherapy-dependent interaction, with high-risk assignment by the 23 gene expression predictor being associated with lower FFS in patients treated with a CHOP/CVP chemotherapy backbone, and better FFS in bendamustine-treated patients.

**Tumor microenvironment**

Since Dave et al. reported that immune response signatures were predictive of patient outcomes, the role of tumor microenvironment has been abundantly studied. Unfortunately, the results from these studies have not led to a coherent understanding of how the immune environment modulates treatment responses in FL. This consideration is likely related to the small size of many studies, the heterogeneous treatments given, the varying methodologies used and also the difficulty to enumerate immune cell subsets in a disease that has an inherent microanatomical organization into follicles and interfollicular regions. Further complicating matters, as many studies use immunohistochemistry, Sander et al. found that concordance for reporting T-cell frequency was low to moderate when performed manually and only moderate to high when using computerized image analysis (54). This may partially explain the heterogeneity of the findings in FL tumor microenvironment studies. As an example of the difficulties to find robust outcome associations with single immunohistochemical markers, we reported that CD163+ tumor-associated macrophages were associated with adverse outcomes in R-CVP-treated patients, whereas the opposite effect was seen in patients from the PRIMA trial who were treated with R-CHOP (55). Many other studies have attempted to associate infiltrating immune cells with patient outcomes, focusing for example on macrophages (40,56-60) or various T-cell subsets or patterns (40,61-68).

Tobin et al. subsequently examined how the tumor microenvironment composition and abundance is related to POD24 (69). They used digital gene expression profiling and found that cases with low immune infiltration and low expression of PD-L2 were enriched for early progression events, which was validated in two independent cohorts. Those cases with low immune infiltration were found to have lower gene expression of immune checkpoint, immune effector and macrophage molecules compared to those...
considered to be high immune infiltration. It currently remains uncertain whether an immune microenvironment-based biomarker using PD-L2 identifies patients at risk of shorter PFS or OS. Going forward, access to large patient cohorts and refined profiling strategies such as single cell measurements of immune phenotypes will increase our understanding of the tumor microenvironment and provide further prognostic insights.

**Functional imaging**

The 2014 Lugano recommendations suggest combined Positron Emission Tomography/Computed Tomography (PET/CT) scan as the imaging modality of choice for initial staging and treatment response assessment for fluorodeoxyglucose-avid lymphomas (70). Total metabolic tumor volume (TMTV) is an effort to standardize quantitative measurement of PET avidity to reflect burden of disease (71). A pooled analysis that was performed on three prospective studies (the PRIMA study, the PET-FOL and the FOLL05 studies) examined the prognostic significance of initial TMTV in patients with FL (23). In total, 185 patients treated with immunochemotherapy were included in the study. Using 510 cm\(^3\) as the cut-off of TMTV, the 5-year PFS was 33% versus 65% and the 5-year OS was 84% versus 95% in patients with high and low TMTV, respectively.

Post-treatment PET positivity has been identified to be a strong independent prognostic risk factor in FL. There have been several large secondary analyses that have looked at the prognostic value of end-of-treatment PET scan. These include the GALLIUM (72), PRIMA (73), GOELAMS (74) and FOLL05 (75). In the PRIMA trial, 26% of patients had PET-positive findings at end-of-treatment (73). The 42-months PFS was 33% and 71% in the PET-positive and PET-negative groups. Similar results were shown in a subset analysis of the FOLL05 trial that was conducted to assess the prognostic value of PET/CT after treatment (75).

A secondary analysis of the GALLIUM trial showed that the percentage of patients who were considered to be in complete response rate by PET/CT compared to CT alone almost doubled (72,75). Twenty-four percent were PET-positive after treatment and PET response at end-of-treatment was prognostic both for PFS and OS. In a prospective study, Dupuis et al. evaluated the impact of PET on patients treated with six cycles of R-CHOP plus two infusions of rituximab (74). Twenty-five percent of patients were PET-positive at end-of-treatment and PET response at end-of-treatment was prognostic both for PFS and OS. In a prospective study, Dupuis et al. evaluated the impact of PET on patients treated with six cycles of R-CHOP plus two infusions of rituximab (74). Twenty-five percent of patients were PET-positive at end-of-treatment and PET response at end-of-treatment was prognostic both for PFS and OS. In a prospective study, Dupuis et al. evaluated the impact of PET on patients treated with six cycles of R-CHOP plus two infusions of rituximab (74). Twenty-five percent of patients were PET-positive at end-of-treatment and PET response at end-of-treatment was prognostic both for PFS and OS. In a prospective study, Dupuis et al. evaluated the impact of PET on patients treated with six cycles of R-CHOP plus two infusions of rituximab (74). Twenty-five percent of patients were PET-positive at end-of-treatment and PET response at end-of-treatment was prognostic both for PFS and OS. In a prospective study, Dupuis et al. evaluated the impact of PET on patients treated with six cycles of R-CHOP plus two infusions of rituximab (74). Twenty-five percent of patients were PET-positive at end-of-treatment and PET response at end-of-treatment was prognostic both for PFS and OS. In a prospective study, Dupuis et al. evaluated the impact of PET on patients treated with six cycles of R-CHOP plus two infusions of rituximab (74). Twenty-five percent of patients were PET-positive at end-of-treatment and PET response at end-of-treatment was prognostic both for PFS and OS. In a prospective study, Dupuis et al. evaluated the impact of PET on patients treated with six cycles of R-CHOP plus two infusions of rituximab (74). Twenty-five percent of patients were PET-positive at end-of-treatment and PET response at end-of-treatment was prognostic both for PFS and OS. In a prospective study, Dupuis et al. evaluated the impact of PET on patients treated with six cycles of R-CHOP plus two infusions of rituximab (74). Twenty-five percent of patients were PET-positive at end-of-treatment and PET response at end-of-treatment was prognostic both for PFS and OS. In a prospective study, Dupuis et al. evaluated the impact of PET on patients treated with six cycles of R-CHOP plus two infusions of rituximab (74). Twenty-five percent of patients were PET-positive at end-of-treatment and PET response at end-of-treatment was prognostic both for PFS and OS. In a prospective study, Dupuis et al. evaluated the impact of PET on patients treated with six cycles of R-CHOP plus two infusions of rituximab (74). Twenty-five percent of patients were PET-positive at end-of-treatment and PET response at end-of-treatment was prognostic both for PFS and OS. In a prospective study, Dupuis et al. evaluated the impact of PET on patients treated with six cycles of R-CHOP plus two infusions of rituximab (74). Twenty-five percent of patients were PET-positive at end-of-treatment and PET response at end-of-treatment was prognostic both for PFS and OS. In a prospective study, Dupuis et al. evaluated the impact of PET on patients treated with six cycles of R-CHOP plus two infusions of rituximab (74). Twenty-five percent of patients were PET-positive at end-of-treatment and PET response at end-of-treatment was prognostic both for PFS and OS. In a prospective study, Dupuis et al. evaluated the impact of PET on patients treated with six cycles of R-CHOP plus two infusions of rituximab (74). Twenty-five percent of patients were PET-positive at end-of-treatment and PET response at end-of-treatment was prognostic both for PFS and OS. In a prospective study, Dupuis et al. evaluated the impact of PET on patients treated with six cycles of R-CHOP plus two infusions of rituximab (74). Twenty-five percent of patients were PET-positive at end-of-treatment and PET response at end-of-treatment was prognostic both for PFS and OS. In a prospective study, Dupuis et al. evaluated the impact of PET on patients treated with six cycles of R-CHOP plus two infusions of rituximab (74). Twenty-five percent of patients were PET-positive at end-of-treatment and PET response at end-of-treatment was prognostic both for PFS and OS. In a prospective study, Dupuis et al. evaluated the impact of PET on patients treated with six cycles of R-CHOP plus two infusions of rituximab (74). Twenty-five percent of patients were PET-positive at end-of-treatment and PET response at end-of-treatment was prognostic both for PFS and OS. In a prospective study, Dupuis et al. evaluated the impact of PET on patients treated with six cycles of R-CHOP plus two infusions of rituximab (74). Twenty-five percent of patients were PET-positive at end-of-treatment and PET response at end-of-treatment was prognostic both for PFS and OS. In a prospective study, Dupuis et al. evaluated the impact of PET on patients treated with six cycles of R-CHOP plus two infusions of rituximab (74). Twenty-five percent of patients were PET-positive at end-of-treatment and PET response at end-of-treatment was prognostic both for PFS and OS. In a prospective study, Dupuis et al. evaluated the impact of PET on patients treated with six cycles of R-CHOP plus two infusions of rituximab (74). Twenty-five percent of patients were PET-positive at end-of-treatment and PET response at end-of-treatment was prognostic both for PFS and OS. In a prospective study, Dupuis et al. evaluated the impact of PET on patients treated with six cycles of R-CHOP plus two infusions of rituximab (74). Twenty-five percent of patients were PET-positive at end-of-treatment and PET response at end-of-treatment was prognostic both for PFS and OS. In a prospective study, Dupuis et al. evaluated the impact of PET on patients treated with six cycles of R-CHOP plus two infusions of rituximab (74). Twenty-five percent of patients were PET-positive at end-of-treatment and PET response at end-of-treatment was prognostic both for PFS and OS. In a prospective study, Dupuis et al. evaluated the impact of PET on patients treated with six cycles of R-CHOP plus two infusions of rituximab (74). Twenty-five percent of patients were PET-positive at end-of-treatment and PET response at end-of-treatment was prognostic both for PFS and OS. In a prospective study, Dupuis et al. evaluated the impact of PET on patients treated with six cycles of R-CHOP plus two infusions of rituximab (74). Twenty-five percent of patients were PET-positive at end-of-treatment and PET response at end-of-treatment was prognostic both for PFS and OS. In a prospective study, Dupuis et al. evaluated the impact of PET on patients treated with six cycles of R-CHOP plus two infusions of rituximab (74). Twenty-five percent of patients were PET-positive at end-of-treatment and PET response at end-of-treatment was prognostic both for PFS and OS. In a prospective study, Dupuis et al. evaluated the impact of PET on patients treated with six cycles of R-CHOP plus two infusions of rituximab (74). Twenty-five percent of patients were PET-positive at end-of-treatment and PET response at end-of-treatment was prognostic both for PFS and OS. In a prospective study, Dupuis et al. evaluated the impact of PET on patients treated with six cycles of R-CHOP plus two infusions of rituximab (74). Twenty-five percent of patients were PET-positive at end-of-treatment and PET response at end-of-treatment was prognostic both for PFS and OS. In a prospective study, Dupuis et al. evaluated the impact of PET on patients treated with six cycles of R-CHOP plus two infusions of rituximab (74). Twenty-five percent of patients were PET-positive at end-of-treatment and PET response at end-of-treatment was prognostic both for PFS and OS. In a prospective study, Dupuis et al. evaluated the impact of PET on patients treated with six cycles of R-CHOP plus two infusions of rituximab (74). Twenty-five percent of patients were PET-positive at end-of-treatment and PET response at end-of-treatment was prognostic both for PFS and OS. In a prospective study, Dupuis et al. evaluated the impact of PET on patients treated with six cycles of R-CHOP plus two infusions of rituximab (74). Twenty-five percent of patients were PET-positive at end-of-treatment and PET response at end-of-treatment was prognostic both for PFS and OS. In a prospective study, Dupuis et al. evaluated the impact of PET on patients treated with six cycles of R-CHOP plus two infusions of rituximab (74). Twenty-five percent of patients were PET-positive at end-of-treatment and PET response at end-of-treatment was prog...
in peripheral blood mononuclear cells (PBMCs), in the bone marrow and/or plasma or serum. High levels of t(14;18) in the PBMCs or bone marrow of FL patients before and after rituximab-chemotherapy induction has been shown to confer worse PFS (75,77-80). Similarly, MRD positivity identifies patients at increased risk of progression in the relapsed/refractory setting after treatment with obinutuzumab and bendamustine, as shown in the GADOLIN trial (81). Quantification of cell-free DNA in plasma can also be used for outcome correlations, as demonstrated by Delfau-Larue et al. (82). The combination of high levels of cell-free DNA and high TMTV appeared to identify a third of patients with a relatively short PFS of 65% at 4 years (82). Sarkozy et al. applied a next-generation sequencing assay that detects lymphoma-specific immunoglobulin rearrangements in plasma samples (83). This method allowed them to determine the abundance of circulating tumor DNA (ctDNA) and 14 patients out of 29 had high levels of ctDNA at diagnosis, which was associated with inferior PFS. While very promising, the application of MRD detection for clinical decision-making has potential shortcomings, especially when relying on t(14;18) quantification. Indeed, the t(14;18) may not be entirely lymphoma-specific as it can be found in peripheral blood samples from healthy individuals (84). Additionally, not all patients are evaluable for MRD, either because their lymphoma is t(14;18)-negative, or because a clonal marker cannot be detected at baseline. The topic of MRD, while intimately associated with progression, is reviewed in more detail elsewhere in this issue.

The detection of gene mutations in ctDNA is gaining momentum. The use of digital droplet PCR has been reported to track hotspot mutations in genes such as EZH2 and STAT6 (85,86). Scherer et al. developed an innovative capture-based sequencing method (named Cancer Personalized Profiling by Deep Sequencing or CAPP-Seq) to detect gene mutations in ctDNA with high accuracy. When applied to serial samples, they identified evolutionary separation, most marked between transformed FL and preceding indolent disease. Moreover, they described a patient who had been treated with rituximab monotherapy for presumed FL, but was subsequently found to have transformation. The mutations that were identified in the pre-rituximab plasma sample classified this sample as transformed FL, suggesting that ctDNA may be a useful approach to detect mutations that are characteristic of aggressive lymphoma, but may be missed in an initial diagnostic tissue specimen, due to intratumoral heterogeneity. The relatively small scale of these studies suggests that ctDNA needs to be further validated before it can be accepted as a definitive prognostic biomarker in FL. This situation is in contrast to DLBCL where early molecular response by ctDNA measurement is increasingly being established as a means to predict early treatment failure (83,87,88). In DLBCL, dynamic risk profiling has been shown to improve outcome prediction, compared to traditional risk models, and could emerge as a platform on which to base decisions to adjust treatment. In FL, the feasibility of detecting gene mutations in plasma is being established, and ctDNA could emerge as a potential biomarker to detect MRD and/or identify patients who are at increased risk of early progression.

**Challenges**

Over the last decade, our understanding of the molecular underpinnings of FL has dramatically increased, and so has the ability to profile large cohorts of samples from trials or other repositories. Consequently, integration of molecular information with established clinical risk factors has been shown to improve risk-stratification prior to initiating immunochemotherapy. Simultaneously, POD24 and end-of-treatment PET scans have emerged as powerful prognostic indicators for patients with FL.

Yet, several challenges remain:

- Molecular testing such as targeted DNA sequencing or digital gene expression profiling is sometimes challenging as many patients get diagnosed with needle core biopsies that may not yield sufficient nucleic acids for genetic assays once the diagnostic workup is complete. Moreover, refined sequencing and gene expression analysis is still largely limited in the up-front clinical setting and is reserved mostly for clinical trials and academic centres, making it rather difficult to apply in most clinical settings.

- FL is a highly heterogeneous disease, clinically and biologically, and a variety of treatment approaches exist to account for the variable clinical presentations. There is certainly also variation around the world in preferred clinical practices. Therefore, a prognostic model that is trained and validated in one group of patients may not be applicable in patients managed with another regimen. For instance, the predominantly clinical models do not appear to have strong prognostic
value in patients treated with rituximab and lenalidomide upfront.

- Beyond inter-patient heterogeneity, many studies have established the universal existence of intra-patient heterogeneity, which provides the substrate for clonal selection and tumor evolution (12,89). Features that are part of risk models may be related to subclones that may not be sampled with any given biopsy for a specific patient. For example, EZH2 mutations can be subclonal, yet have important prognostic information (90). Moreover, striking clonal divergence underlies transformation to aggressive lymphoma, potentially contributing to the difficulty to predict transformation (44,89).

- POD24 as a concept, and end-of-treatment PET imaging cannot be applied before the initiation of frontline therapy. Hence, they can—by definition—not be applied to decide on an alternative treatment strategy.

- There is no currently accepted method to intensify treatment in patients identified as having high-risk disease by any prognostic model. Intensifying treatment with autologous stem cell transplant does not improve outcome in patients with high-risk m7-FLIPI or POD24-PI (91). In addition, the FLIPI, m7-FLIPI and POD24-PI classify 44%, 21% and 33%, respectively of patients into the high-risk category, despite that they do not progress within 24 months, which could lead to overtreatment if intensification of treatment was available (92).

**Outlook**

Beyond current challenges, our enhanced understanding of lymphoma biology and the increasing availability of novel therapeutic approaches lead to opportunities that will likely allow us to refine prognostication and select biology-informed treatments (Figure 2). Very recently, studies have been presented that describe how various chemotherapy backbones influence patient outcome based on the presence of specific gene alterations such as EZH2 mutations. While these studies need to be validated, they raise the possibility that relatively limited genetic tests could help us select the most suitable frontline standard-of-care regimen. Moreover, novel therapies such as epigenetic inhibitors, signaling inhibitors and immune therapies are increasingly studied in the relapsed/refractory setting and will likely have value in the frontline setting. Correlative studies will determine their efficacy based on the specific molecular landscape and

---

**Figure 2** Schematic overview of prognostic tools that can be used to predict which patients experience early progression (figure created with BioRender.com).
immune contexture in a given patient’s tumor.

Treatment could also be adapted by escalation or de-escalation based on dynamic measures of response during initial first-line treatment, which is already routine in other lymphoma subtypes such as Hodgkin lymphoma (93,94). Although interim PET/CT during first-line treatment has not been demonstrated to be a robust marker of outcome (34), another approach that is currently being assessed is the use of ctDNA which is described above. Clinical trials need to ascertain the clinical utility of changing therapy based on suboptimal response and high risk of early progression, for example by switching to orthogonal therapies such as those that leverage the immune system. Lastly, ascertainment of the depth of response at the end of induction may guide optimal strategies for maintenance. The time has come to leverage the availability of robust biomarker platforms to translate our increasing understanding of lymphoma biology into improved patient outcomes.

Acknowledgments

Funding: This work was supported by a Terry Fox Research Institute New Investigator Award, a Leukemia & Lymphoma Society of Canada Operating Grant, a Leukemia & Lymphoma Translational Research Program and the Princess Margaret Cancer Foundation (support to RK). AS is supported by a Postdoctoral Fellowship from the Canadian Institutes of Health Research.

Footnote

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/aol-20-46). The series was commissioned by the editorial office without any funding or sponsorship. RK reports grant funding from Gilead Sciences and Roche, and research funding from AstraZeneca. The authors have no other conflicts of interest to declare.

Provenance and Peer Review: This article was commissioned by the Guest Editors (Mark Roschewski, Carla Casulo) for the series “Follicular Lymphoma” published in Annals of Lymphoma. The article has undergone external peer review.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References


Cite this article as: Liu Q, Silva A, Kridel R. Predicting early progression in follicular lymphoma. Ann Lymphoma 2020.

doi: 10.21037/aol-20-46