2012

Risk factors of follicular lymphoma

Shuangge Ma, *Yale University*

Risk Factors of Follicular Lymphoma

Shuangge Ma

Yale University

Introduction: Non-Hodgkin Lymphoma (NHL) is a heterogeneous group of malignancies with over thirty different subtypes. Follicular lymphoma (FL) is the second most common subtype and the most indolent one. It has morphologic, immunophenotypic and clinical features significantly different from other subtypes. Considerable effort has been devoted to the identification of risk factors for the etiology and prognosis of FL. Those risk factors may potentially advance our understanding of the biology of FL and more importantly have an impact on clinical practice.

Areas covered: We first very briefly review the epidemiology of NHL and FL. For FL etiology and prognosis separately, we review the clinical, environmental and omics (genetic, genomic, epigenetic …) risk factors suggested in the literature.

Expert opinion: A large number of potential risk factors have been identified in recent studies. However, there is a lack of consensus, and many of the suggested risk factors have not been rigorously validated in independent studies. There is a need for large-scale, prospective studies to consolidate existing findings and discover new risk factors. Some of the identified risk factors are successful at the population level. More effective individual-level risk factors/models remain to be identified.

Keywords: Follicular lymphoma; Etiology; Non-Hodgkin lymphoma; Prognosis; Risk factor.

1. Introduction

Non-Hodgkin Lymphoma (NHL) is a heterogeneous group of malignancies of lymphocyte origin. NHL usually arises or is present in lymphoid tissues, such as lymph nodes, spleen and bone marrow. During the past three decades, there have been consistent reports of increase in the incidence of NHL worldwide. In general, age-adjusted incidence rates of NHL are higher in more developed countries. In the United States, the age-adjusted incidence rate has almost doubled since the 1970s from 11.07/100,000 in 1975 to 20.20/100,000 in 2008 [1]. It is the fifth most commonly diagnosed malignancy in the US among both men and women. According to the National Cancer Institute, it is estimated that 66,360 new cases of NHL were diagnosed in 2011, with 19,320 deaths. The Lymphoma Research Foundation estimates that 332,000 Americans are currently living with NHL. There are over thirty NHL subtypes, with the two most common subtypes – diffuse large B-cell lymphoma (DLBCL) and follicular lymphoma (FL) – accounting for about 30% and 20% of all NHL cases, respectively [2].

FL is defined as a lymphoma of follicle center B-cells, which has at least a partially follicular pattern [2]. It is positive for the B-cell markers CD10, CD19, CD20 and CD22, but almost always negative for CD5. It is the most common form of indolent NHL. According to the WHO criteria, FL can be morphologically graded into grade 1 (<5

centroblasts per high-power field (hpf)), grade 2 (6-15 centroblasts/hpf) and grade 3 (>15 centrobalsts/hpf). Grade 3 can be further subdivided into grade 3A (centrocytes still present) and grade 3B (the follicles consist almost entirely of centroblasts). Grades 1, 2, and 3A are considered to be indolent and incurable, whereas grade 3B is considered an aggressive but curable disease similar to DLBCL. It has been noted that although this grading system is valuable from a pathological perspective, its clinical relevance is still debatable. Over time, histologic transformation of FL from an indolent disease to a DLBCL may occur in 10-70% of patients, with an estimated risk of 3% per year, and is associated with rapid progression of lymphadenopathy, extranodal disease, B symptoms and elevated serum LDH [3]. FL may also transform to Burkitt lymphoma or other types of aggressive lymphomas, although much less commonly.

Compared with some other forms of NHL, for example DLBCL, FL usually progresses slowly. Despite the fact that most FLs are advanced at the time of diagnosis, the median survival of patients with FL is approximately 8-10 years, and many patients may not require treatment for a long time. Several retrospective studies have shown an important improvement in overall survival of FL patients in the last fifteen years when compared to historical controls. The improvement has been largely attributed to the introduction of anti-CD20 monoclonal antibodies (MoAba) in the treatment [4].

Considerable effort has been devoted to the identification of NHL risk factors [5]. In this article, we focus on FL and refer to other publications for discussions on other types of NHL. “Classic” research on FL risk factors has been focused on clinical measurements and environmental exposures. More recently, with the fast development of high-throughput profiling techniques, genome-wide analysis has been extensively

conducting using various platforms including gene expression profiling, array comparative genomic hybridization (aCGH), single nucleotide polymorphisms (SNP) arrays and several other novel technologies that measure methylation status and epigenome. In the following sections, we review risk factors identified for FL etiology and prognosis separately. For each aspect, clinical measurements, environmental exposures, and omics risk factors are reviewed. In addition, we also provide brief discussions on several pitfalls in the pursuit of FL risk factors. We acknowledge that the identification of FL risk factors is an extremely complex process, involving a large number of steps including study design, execution, analysis, validation and others. Due to the limited scope of this article, inevitably, some important aspects will be missed. We refer to [5,6,7] and references therein for related discussions.

2. Etiology

2.1 Clinical and environmental risk factors

Overall, the etiology of NHL is poorly understood [5]. It has been suggested that age, gender, and ethnicity may affect a person's likelihood of developing FL. The incidence of FL increases with age. Although in principal FL may occur at any age, it is extremely rare in children. The median age at diagnosis is 60-65 years. Women have a slightly higher risk of developing FL than men. The incidence of FL is low among Chinese and Japanese. People of Jewish ancestry have a higher incidence of lymphoma. In the US, the incidence is 2-3 times higher in Caucasian than in African-American.

Risk factors that may increase the risk of NHL also include medications that suppress the immune system (for example if a person has just had an organ transplant, he/she is more susceptible because immunosuppressive therapy has reduced the body's ability to fight off new illnesses), and infection with certain viruses and bacteria. Viruses that have been implicated in the development of FL include the Epstein-Barr virus (EBV), human T-cell lymphotropic virus (HTLV) type I, and the herpesvirus associated with Kaposi sarcoma (i.e., human herpesvirus HHV-8) [8]. It is worth pointing out that although they have been implicated in the development of FL, these viruses are linked mostly with diffuse or high-grade lymphomas. Congenital immunodeficiencies have been associated with lymphoma. Acquired immunodeficiencies may include infection with the human immunodeficiency virus (HIV). Note that most lymphomas associated with HIV are intermediate-grade or high-grade lymphomas.

Certain chemicals, such as those used to kill insects and weeds, may increase the risk of developing NHL. An increased risk of FL was found among women who started using hair dyes before 1980, and cannot be excluded for women who started using in 1980 and after [9]. However, more research is needed to better understand the mechanical link between pesticides and the development of FL.

Multiple lifestyle factors may also contribute to the risk of FL. However, it is noted that some conflicting results have been reported in the literature. A case-control study conducted in Italy has linked tobacco use to the development of FL [10]. The researchers reported a 50% increased risk for 16-33 pack-years to an 80% increased risk with 34 pack years or greater. There was no increased risk for the other NHL subtypes. In a second study, analysis of data on over 6,500 NHL cases also suggested

a positive link between smoking and incidence of FL but not other NHL subtypes. United States researchers from Yale University have reported that long-term cigarette smoking in women increased the risk of developing FL [13]. Analysis of the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial, which had 1,264 histologically confirmed NHL cases, suggested that smoking was not associated with NHL overall but was inversely associated with FL (ever smoking vs never: hazard ratio(HR)=0.62, 95% confidence interval [0.45, 0.85]) [12]. Alcohol intake may directly affect immune function, which is an important etiologic factor for lymphoid malignancies. Morton and others [13] conducted a pooled analysis of nine case-control studies of NHL and showed that ever drinkers, compared with never drinkers, had a 17% lower risk of NHL. This finding can be potentially explained by a beneficial effect of moderate alcohol consumption on immune system. However, no significant dose-response relationship was observed, which could be used to argue against a possible biological mechanism for the observed inverse association. Analysis of the PLCO data suggested that alcohol consumption was unrelated to NHL (drinks/week: p-value for trend=0.187) [12]. In six large prospective cohort studies of alcohol and NHL, researchers found that moderate alcohol intake was associated with a 41% reduction in risk among Iowa women, and heavier alcohol intake was associated with a 33% risk reduction among United Kingdom women, a 23% risk reduction among retired US men and women, and a 40% reduction among Japanese men. However, alcohol intake was not associated with NHL risk among Finnish male smokers or US men and women in a cancer screening trial. Chang and others [14] investigated whether the history of alcohol drinking affected the risk of NHL using the California Teachers Study cohort, a prospective cohort with 496 women

diagnosed with B-cell NHL. It was found that women who were former alcohol drinkers at baseline were at an elevated risk of overall B-cell NHL and FL (rate ratio=1.81, 95% CI [1.00, 3.28]). The researchers argued that it was important to identify both current and past alcohol consumption status. In the Netherlands Cohort Study, it was found that the rate ratio of lymphatic malignancies per 4-unit increase in BMI (body mass index) at 20 years of age was 1.13 (95% CI [1.01, 1.25]). The overall rate ratio of lymphatic malignancies per 5-cm increase in height was 1.08 (95% CI [1.02, 1.15]). The rate ratio for FL alone was not significant and at 1.15 (95% CI [0.95, 1.40]) [15]. In a prospective cohort study with 37,931 Iowa women among whom there were 261 cases of NHL and 58 cases of FL, Cerhan and others [16] observed no overall association between anthropometric characteristics (including BMI) and risk of overall NHL or FL. In contrast, in a population-based case-control study, Skibola and others [17] found that the risks of NHL and FL were positively associated with an overweight or obese status compared to a normal-weight status. In a Scandinavian study with 3,055 NHL cases and 3,187 population-based controls, Cheng and others did not find any association between BMI and risk of FL [18].

2.2 Omics risk factors

As with other types of cancers and other subtypes of NHL, there is increasing evidence that omics (genetic, genomic, epigenetic…) risk factors may have independent contribution to the risk of FL beyond clinical risk factors and environmental exposures [5,19,20,21].

The most common acquired nonrandom chromosomal translocation in FL patients is the t(14; 18) translocation, which is found in more than 80% of all cases. This translocation results in the overexpression of the BCL-2 gene, which encodes apoptosis regulator proteins and has been implicated in a number of cancers including melanoma, breast, prostate and lung carcinomas. In fact, the detection of the t(14, 18) product by polymerase chain reaction (PCR) has been frequently used in the diagnosis and follow-up of patients with FL. In some cases, the presence of BCL-2 staining in biopsies may be significant for the patient's prognosis or likelihood of relapse. However, this translocation is not unique to FL. It has been detected in healthy patients as well as patients with other types of tumors. A small fraction of FL (~5%) does not exhibit the classical t(14;18) but instead contains alterations affecting BCL-6 at 3q27, including t(3;14)(q27;q32) [22]. This leads to deregulated expression of the transcriptional repressor BCL-6, which is normally required for germinal centers formation. Protein encoded by gene BCL-6 acts as a sequence-specific repressor of transcription and has been shown to modulate the STAT-dependent Interleukin 4 (IL-4) responses of B cells.

High-throughput microarray studies have been conducted, searching for genes whose expressions are associated with the etiology of FL [23]. Gene expression profiling of normal germinal center B (GCB) cells has been shown to be unchanged in FL, supporting the perspective that FL arises from this stage of B-cell differentiation. In a cDNA microarray study with 588 genes, Husson and others [24] identified 28 genes that were down-regulated and 37 genes that were up-regulated in FL cells compared with normal GCB cells. The expression level of each differentially expressed gene was then verified by quantitative PCR, resulting in 24 up-regulated genes and 8 down-

regulated genes (with p-value<0.10). Up-regulated genes in FL included 2 cell cycle regulator proteins that are involved in G1 arrest, p21 and p16, which is consistent with the low proliferative nature of FL cells. The identified up-regulated genes also included cell cycle regulator proteins (CDK10, p120, p21, and p16), genes that are involved in cell-cell interactions (TNF, IL2RG, and IL4RA), and the transcription factors PAX5 and Id-2 that are involved in norm B-cell development. Down-regulated genes in FL included MRP8 and MRP14, which are involved in adhesion.

Conde and others [25] conducted a three-stage genome-wide association study and identified two variants associated with FL at 6q21.32 (rs10484561 and rs7755224). The combined sample contained 1,465 FL cases and 6,011 controls. For rs10484561, the combined OR (odds ratio) was 1.95 (95% CI [1.72, 2.22]) and the combined p-value=1.12×10^{-29}. This marker is in the human leukocyte antigen (HLA) class II region. Smedby and others [26] conducted a three-stage genome-wide association study, starting with a genome-wide scan of 379 FL cases and 791 controls followed by validation in 1,049 cases and 5,790 controls, and identified a second independent FL-associated locus on 6q21.32, rs2647012 (combined OR=0.64, combined p-value=2×10^{-21}), located 962 base pairs away from rs10484561 ($r^2 < 0.1$ in controls). After mutual adjustment, the associations at the two SNPs remained genome-wide significant (rs2647012: OR=0.7, adjusted p-value=4×10^{-12}; rs10484561, adjusted OR=1.64, adjusted p-value=5×10^{-15}). Haplotype and coalescene analyses indicated that rs2647012 arose on an evolutionarily distinct haplotype from that of rs10484561 and tagged a novel allele with an opposite (protective) effect on FL risk.

Epigenetic events have important implications in cancer. Recent molecular studies have established that activation of various oncogenes and silencing of tumor suppressor genes is required for FL development and progression [27]. Commonly methylated genes in FL include androgen receptor gene, which encodes a member of the receptor group that binds and mediates the actions of androgens. DNA promoter hypermethylation of the androgen receptor gene is common in FL. McDonald and others [28] reported 16/19 samples of T and B lymphomas positive for methylated androgen receptor genes. A similar analysis of grade 1 and 2 FL found 25/26 samples tested were positive for methylation at the androgen receptor promoter region [29]. SHP1 is a phosphotyrosine phosphatase with important roles in the regulation of immune system cell differentiation and activation. Methylation of SHP1 promoter region appears common across multiple subtypes of lymphomas (including DLBCL, MALT lymphoma, plasmacytomas and mantel cell lymphoma). In [30], 32/33 FL samples assayed were methylated at the SHP1 promoter region. In contrast, no promoter methylation of SHP1 was observed in 20 normal reactive lymph node samples. The almost uniform methylation of SHP1 in FL samples and the high rates across several lymphoma subtypes suggest that the down-regulation of SHP1 due to promoter hypermethylation may play an important role in lymphomagenesis. DAPK (dead-associated protein kinase) is a calcium-calmodulin-dependent serine/threonine kinase that is involved in apoptosis. Methylation of DAPK may be common in FL. In two published studies [31,32], 25/29 FL samples were positive for aberrant DAPK methylation. The finding is consistent with the hypothesis that methylation of DAPK is a common early epigenetic phenomenon of FL, which allows cells to escape the normal

apoptosis process. p16 (cyclin-dependent kinase inhibitor 2A) is located on chromosome 9q21 and the most commonly altered gene in human malignancies. Homozygous deletion of p16 is common in multiple cancers, however, not in patients with newly diagnosed FL. Two published [33,34] have investigated p16 deletions in FL. A combined analysis of published data showed aberrant p16 promoter methylation in 5/16 FL cases. p15 is also a cyclin-dependent kinase inhibitor and located 25 kilobases from the p16 gene. Its protein shares significant areas of homologous amino acid sequences with p16. Methylation of the p15 promoter region has been reported in acute myeloid and lymphoblastic leukemia, multiple myeloma, and lymphomas. However, few studies have specifically detailed the histologic subtype of the lymphomas analyzed. Of those that have, 10/27 FL samples were found to have detectable p15 methylation [35,36]. The p57 gene (cyclin-dependent kinase inhibitor 1C) is located on chromosome 11p15.5. The expression of p57 gene is absent in various hematological cell lines. In [37], 8/18 FL samples were observed to have methylation of the p57 promoter region. p14 protein is also a cyclin-dependent kinase inhibitor known to induce cell cycle arrest at G1 and G2. Mice with homozygous deletion of p14 frequently develop numerous tumors including lymphomas. Despite interesting findings in animal studies, Baur and others [35] failed to find methylation of the p14 promoter region in FL samples. Hypermethylation of p14 does not appear to be a common epigenetic event in FL. As suggested in [27], genes whose methylation may be involved in FL etiology also include GSTP1, IL-12 receptor beta-2, Snk/plk2 and GADD45-gamma. More research is needed to draw definitive conclusion on these genes.

3. Prognosis

3.1 FLIPI and FLIPI2

The prognosis of FL has attracted extensive attention. For a brief review of the history on FL prognosis research, we refer to [38] and references therein.

In 2004, the Follicular Lymphoma International Prognostic Index (FLIPI) was built from a retrospective analysis of more than 4,000 patients with FL treated between 1985 and 1992 [38]. The endpoint was OS (overall survival). In a multivariate analysis, five parameters were used to build this index: age (>60 versus ≤60), serum LDH level (>upper limit of normal (UPLN) versus ≤UPLN), number of nodal areas (>4 versus ≤4), hemoglobin level (<120 versus ≥120g/L) and Ann Arbor state (III-IV versus I-II). Three risk groups, low, intermediate and high, were distinguished. The low risk group is defined as “number of risk factors=0-1”. It accounts for 36% of the patients, with five-year OS 90.6% and ten-year OS 70.7%. The intermediate group is defined as “number of risk factors=2” and accounts for 37% of the patients, with five-year OS 77.6% and ten-year OS 50.9%. The high risk group contains the rest 27% of the patients, with five-year OS 52.5% and ten-year OS 35.5%. Using independent cohorts of patients, several retrospective studies have confirmed the effectiveness of the FLIPI. Particularly, the distributions of patients and differences in hazard ratios for death between different prognostic subgroups were found to be very similar to those in [38]. The FLIPI has thus been considered as a standard prognostic index for FL. In addition, although the FLIPI was originally constructed for OS, in some recent studies, it has been used as a prognostic index for PFS (progression-free survival). FLIPI may have several limitations.

For example, even though it has been used in some PFS studies, the FLIPI was not originally designed for PFS. In addition, the samples were collected between 1985 and 1992, during which period some important clinical and biological parameters were not collected and thus not used to construct FLIPI. In addition, patients in the FLIPI cohort only received conventional chemotherapy or no treatment, making them significantly different from those in more recent studies.

The FLIPI2 was built to solve the problems encountered by FLIPI [4,39,40]. It has PFS as the endpoint, as opposed to OS with FLIPI, and integrates all the modern parameters prospectively collected. It also has five prognostic parameters: longest diameter of the largest tumor mass >6 versus ≤ 6cm, serum β2 microglobulin level >UPLN versus ≤UPLN, bone marrow involved or not, hemoglobin level ≤120 versus >120 g/L, and age >60 versus ≤60 years. In the analysis of 812 patients, 88% of whom were treated with rituximab, FLIPI2 classified patients into three risk groups. The low risk group was defined as having none of the risk factors and accounted for 20% of the patients. The 3-year PFS was 91%, and the 5-year PFS was 79.5%. The intermediate group had 1-2 of the risk factors and accounted for 53% of the patients. The 3-year PFS was 69%, and the 5-year PFS was 51%. The high risk group had 3-year PFS 51% and 5-year PFS 19%. Arcaini and others [41] conducted retrospective analysis of 280 patients, among whom 262 were diagnosed after 1995 and 190 were treated with Rituximab, and confirmed the accuracy of FLIPI2 for PFS. Compared with FLIPI, more independent confirmation studies on FLIPI2 are needed. Note that FLIPI2 has an endpoint different from that of FLIPI and is not intended to replace FLIPI. Instead, they complement each other.

3.2 Other clinical and environmental risk factors

Lifestyle factors also have an impact on prognosis. Geyer and others [42] evaluated the association between pre-diagnosis cigarette smoking, alcohol use, and BMI with overall survival in 1,286 patients enrolled through population-based registries in the US from 1998 through 2000. It was found that compared with never smokers, former smokers (HR=1.59, 95% CI [1.12, 2.26]) and current smokers (HR=1.50, 95% CI [0.97, 2.29]) had poorer survival, and poorer survival was found to be positively associated with smoking duration, number of cigarette smoked per day, pack-years of smoking, and shorter time since quitting (all p-value<0.01). Alcohol use was associated with poorer survival (p=0.03) compared with nonusers. Those who drank >43.1 g/week (median intake among drinkers) had poorer survival (HR=1.55, 95% CI [1.06, 2.27]), whereas those drinkers consuming less than this amount demonstrated no significantly different survival (HR=1.13, 95% CI [0.75, 1.71]). Greater BMI was associated with poorer survival (p=0.046), but this survival disadvantage was only observed for obese individuals (HR=1.32 for BMI≥30 versus BMI 20-24.9; 95% CI [1.02, 1.70]). Han and others [43] were among the first to test the hypothesis that a higher intake of fruits and vegetables is associated with better NHL survival [44]. Using a population-based cohort of 568 women with newly diagnosed NHL followed for a median of 7.7 years, the researchers found 32% better overall survival for NHL patients who had higher pre-diagnosis intake of vegetables and fruits (HR=0.68, 95% CI [0.49, 0.95]) after adjustment for demographic and clinical variables. Total vegetables (HR=0.58, 95% CI [0.38, 0.89]), green leafy vegetables (HR=0.71, 95% CI [0.51, 0.98]) and citrus fruits

(HR=0.73, 95% CI [0.54, 0.99]) showed the strongest associations. It is interesting that all these associations held for both FL and DLBCL.

3.3 Omics risk factors

Dave and others [45] conducted microarray gene expression profiling of 191 FL patients, searching for genes associated with clinical prognosis. The researchers defined two expression profiles, referred to as the immune-response-1 and -2 signatures, associated with long and short survival, respectively. The immune-response 1 signature consisted of T-cell-specific genes CD7, CD8B1, LEF1, ITK and STAT4 and macrophage lineage genes ACTN1 and TNFSF13B (BAFF). The immune-response-2 signature included genes expressed by macrophages and/or dendritic cells, such as TLR5, FCGR1A, SEPT10, LGMN and C3AR1 (complement 3a receptor 1). Cell sorting experiments confirmed that the immune-response signatures mainly reflected expression levels of the various, non-neoplastic CD19- cell populations. It is noted that, Tibshirani [46] analyzed the same data using various standard statistical techniques and argued that "… our analysis sheds serious doubt on the reproducibility of the authors' (Dave et al.) biologic findings", highlighting the analytic challenges faced by high-throughput gene expression studies [5]. Several research groups compared gene expression profiles of low-grade FLs (histological grades 1–2) with those of grade 3 FLs or FLs that had transformed into DLBCLs. Glas and others [47] conducted supervised classification on a training set of paired samples from patients who experienced either an indolent or aggressive disease course, and established a gene expression profile

with 81 genes. It was shown that in the training set, this gene signature had 100% accuracy distinguishing low-grade from high-grade diseases. The classification accuracy was 93% in an independent validation set. In a third set of FL samples where histologic grading was ambiguous, this gene signature showed a classification accuracy of 94%. Genes that were significantly up-regulated in the aggressive phase of the disease included those involved in cell cycle control (such as genes CCNE2, CCNA2, CDK2, CHEK1 and MCM7) and DNA synthesis (including genes TOP2A, POLD3A, HMGA1, POLE2, GMPS and CTPS) as well as those reflecting increased metabolism (including genes FRSB, RARS, HK2 and LDHA) and activation of several signaling pathways (including genes FRZB, HCFCR1, PIK4CA, and MAPK1). Genes that are derived from the reactive infiltrate of T cells and macrophages (CD3C, CXCL12 and TM4SF2) were up-regulated in the indolent phase of the disease. Lossos and others [48] profiled 12 FLs with transformation and identified a set of 671 genes that exhibited at least a threefold variation in the biopsy pairs of three or more patients. The researchers identified two distinct gene profiles possibly associated with FL prognosis. Five out of the 12 cases displayed enhanced expression of C-MYC and its target genes, whereas in four cases a decreased expression of C-MYC and its target genes was observed. De Vos and others [49] profiled four FLs with documented progressions. Among the top 36 up-regulated and 36 down-regulated genes, seven genes were also identified by [48]. Genes CDA and GAPD, two genes reflecting levels of metabolism, were among the overlapped and up-regulated. Genes IRF8 and of PTPRC were also identified in the two studies and down-regulated. The researchers also noted down-regulation of different T-cell markers upon transformation such as CD7, FYB (Fyn

binding protein) and SEMA4D (CD100). Elenitoba-Johnson and others [50] studied 11 FLs that transformed into DLBCL. The findings included 67 significantly up-regulated and 46 down-regulated genes in DLBCL. Interesting up-regulated genes included the growth factor/cytokine receptors MET (the hepatocyte growth factor receptor), FGFR3 (fibroblast growth factor receptor 3), LTBR (lymphotoxin b receptor) and PDGFRB (platelet-derived growth factor receptor b). In addition, gene p38BMAPK was also found up-regulated in DLBCL. This finding was confirmed in follow-up mechanical study. Janikova and others [51] conducted gene expression profiling of 31 non-selected patients with FL, 12 of whom were in relapse and the remaining 19 newly diagnosed. The researchers employed template matching and defined two gene sets composed of genes differentially expressed among samples. These gene sets shared an over-representation of genes with similar biological functions and were termed T-CELL and PROLIFERATION profiles. The poor profile was defined by a high PROLIFERATION score and/or low T-CELL score. The poor profile cohort contained a significantly higher proportion of relapsed cases. In addition, a comparison of samples from initial diagnosis and from relapse showed significant differences mainly in the T-CELL profile.

With the analysis of 278 patients, Cerhan and others [52] found that SNPs in genes IL8 (rs4073; HR_{TT}=2.14, 95% CI [1.26, 3.63]), IL2 (rs2069762, $HR_{GG/TT}$=1.80, 95% CI [1.06, 3.05]), IL12B (rs3212227; $HR_{AC/CC}$=1.83, 95% CI [1.06, 3.06]) and IL1RN (rs454078; HR_{AA}=1.93, 95% CI [1.11, 3.34]) were the most significant predictors of survival. A summary score using the number of deleterious genotypes from these genes was shown to be significantly associated with survival (p=0.001). A risk score that combined the four SNPs with clinical and environmental risk factors was even more

strongly associated with survival ($p<0.001$). The 5-year survival estimates were 96%, 72% and 58% for the low, intermediate and high risk groups, respectively. Wrench and others [53] confirmed the association of SNP rs10484561 and rs6457327 with risk of FL and demonstrated that SNP rs6457327 predicted both time to and risk of FL transformation independently of clinical and environmental risk factors, particularly including FLIPI. Han and others [54] investigated overall survival of 101 FL patients. This study took a candidate-gene approach. A total of 1,229 SNPs representing 122 KEGG pathways were analyzed. The researchers identified two pathways as having independent predictive power for prognosis beyond clinical and environmental risk factors. The two pathways were Endometrial cancer, which included ten SNPs from genes CASP9, CCND1, MLH1, MYC, TP53 and CTNNB1, and pathway Melanogenesis, which included five SNPs from genes MC1R and CTNNB1. It was noted that the identified pathways differed significantly from those for DLBCL and NHL overall. With the same data, Ma and others [55] adopted a novel data mining approach using thresholding and identified 131 genes (187 SNPs) as associated with prognosis of FL. The identified genes/SNPs represented the following KEGG pathways: apoptosis (including genes CASP9, IL1A, IL3, IRAK2, IRAK3), Cytokine-cytokine receptor interaction (gene IFNGR2, IL10, IL15, IL1A, IL3, IL4R), focal adhesion (RAC1, RAC2), Leukocyte transendothelial migration, and multiple signaling pathways (including B cell receptor signaling pathway, calcium signaling pathway, Fc epsilon RI signaling pathway, Insulin signaling pathway, Jak-STAT signaling pathway, MAPK signaling pathway, p53 signaling pathway, Toll-like receptor signaling pathway, VEGF signaling pathway, and Wnt signaling pathway).

Studies have linked 1p36 deletions with worse prognosis in FL [56]. With a cohort consisting of 251 specimens, the researchers identified 46 cases with nonsynonymous mutations affecting TNFRSF14 (tumor necrosis factor receptor superfamily member 14). OS and disease-specific survival (DSS) were associated with the present of TNFRSF14 mutation in patients whose overall treatment included rituximab. It was further shown that inferior OS and DSS were most pronounced in patients whose lymphomas contained both TNFRSF14 mutations and 1p36 deletions after adjustment for FLIPI.

O'Shea and others [57] studied 185 FL patients to assess the prognostic relevance of aUPD (acquired homozygosity in the form of segmental acquired uniparental disomy). It was found that the number of genetic abnormalities was predictive of outcome. In particular, the presence of more than three abnormalities was associated with inferior OS. Sites of recurrent aUPD were detected on 6p, 16p, 12q, 1p36, 10q and 6q. In multivariate analysis, aUPD on 1p36 correlated with shorter OS. aUPD on 16p was predictive of transformation and correlated with poorer PFS.

With its important implications in the etiology of FL, methylation was also investigated as potential risk factors for prognosis [58]. It was found that methylation profiles were conserved in sequential FL and transformed-FL biopsies, suggesting that widespread methylation represents an early event in lymphomagenesis and may not contribute substantially to FL transformation. Significant ($p<0.05$) correlation between FL methylation values and reduced gene expression was demonstrated for up to 28% of loci. Such findings suggested that it may not be sensible to attribute FL prognosis to methylation.

4. Discussion

Identification of FL risk factors is of significant interest. In cancer study overall, it is estimated that one-third of cancer occurrences are preventable, and that one-third of people with cancer can recover fully if the cancer is detected at an early stage. Identifying risk factors contributing to FL etiology can help specify high risk population and facilitate preventive actions. FL is an indolent disease. For many patients, a “watch and wait” strategy may be sufficient [59,60]. Thus, it is important to single out patients possibly with an aggressive path for active treatment. Some recent studies such as [5] investigate NHL overall. Different subtypes of NHL, although sharing a certain degree of similarity, differ significantly from each other. From a practical point of view, it is necessary to study each subtype separately.

For presentation clarity, we provide separate discussions on etiology and prognosis, and on clinical/environmental and omics risk factors. The development and progression of FL is a continuous process and caused by the combined effects of multiple types of risk factors. In this article, we review a large number of potential risk factors suggested in the literature. We note that most of them have not been extensively validated in independent studies, and there is thus a lack of consensus. Even the successful ones, such as the FLIPI, may be of limited practical value. A few possible reasons for the lack of success are provided in the “Expert Opinion” section. Research on FL is moving fast. In addition, because of our limited knowledge in this area, our

review is far from comprehensive. The present review may need to be updated in the near future.

Identification of risk factors is the first step in clinical practice. Even though a large number of possible risk factors have been suggested, most of them have not been effectively utilized in clinical practice, particularly personalized risk stratification and treatment selection. Due to the limited scope of this article, we refer to [3,4,61,62] for discussions on FL treatment regimens and implications of identified risk factors in initial treatment selection.

5. Expert opinion

In the literature, a large number of possible FL risk factors have been suggested, some of which are reviewed in this article. Unfortunately, many of them have low reproducibility and have not been successfully validated in independent studies. As discussed in [5,63] and references therein, multiple factors may contribute to the low reproducibility. The study cohorts used in different studies may have significantly different clinical and demographic characteristics, casting serious doubt on the comparability of risk factors identified in those studies. In addition, patients in different studies may have different treatment regimens. In prognosis study, the treatment regimen has a tremendous impact on the prognostic path and should be accounted for in the prognosis modeling and risk factor identification.

Compared with clinical and environmental risk factors, omics data are considerably more complicated, and so the identification of omics risk factors can be

much more challenging. There are several pitfalls in omics studies. For example, different studies adopt different technologies (gene expression profiling, IHC, flow cytometry), which do not measure the same features. Gene expression studies make the inaccurate assumption that gene and protein expression are perfectly matched. IHC studies often fail to report on immunoarchitectural features (the distribution of cells in relation to the malignant follicles) that may be more important than the total number of cells. Flow cytometry techniques offer the possibility of more objectively enumerating many thousands of cells, but interpretation of these data is very much operator-dependent and similarly does not address aspects of immunoarchitecture. In addition, as discussed in a recent study [5], existing analytic techniques for omics data have limitations, and there is a need for more effective analysis methods.

As with other subtypes of NHL and other cancers, the development and progression of FL is a complex process and involves the interactions of multiple clinical risk factors, environmental exposures, genetic, genomic, epigenetic events and others. Most existing studies and analysis approaches only target a subset of potential risk factors, which may result in biased estimates and false risk factor identification. For some cancers, there are studies attempting to systematically investigate and integrate a large number of different types of risk factors. However, to the best of our knowledge, there is still no such study on FL.

In several prognosis studies, the FLIPI has been shown to be effective. Similar success has been observed for a few other risk factors/indexes. However, it has been noted that these risk factors are only effective at a population level, predicting “average” risk and progression for a group of subjects. The prediction accuracy for a single subject

has been disappointingly low. In recent cancer research, a lot of attention has been devoted to personalized medicine, with the ultimate goal of predicting cancer development and progression at an individual level. For FL, there is insufficient attention on personalized predictive modeling and risk factor identification.

Our recommendations for future research include the following. First, existing literature needs to be more carefully and more systematically examined. Risk factors suggested in multiple studies warrant a higher priority in future validation studies. In addition, it is of significant interest to examine the causes of discrepancies (in the identified risk factors). The differences between cohorts' characteristics and technical aspects (such as platforms and profiling protocols in omics studies) need to be quantified. Such examination may help determine how much discrepancy can be attributed to heterogeneity among studies and improve reproducibility of identified risk factors. Second, it is worthwhile to reanalyze existing studies, particularly omics studies, using more advanced analytic methods. For example, in most published gene expression and SNP studies, the interplay among genes is not sufficiently accounted for. There is a need to reanalyze such data using methods that take a system biology point of view. We refer to [5] for more detailed discussions. Third, there is a need for carefully-designed, prospective studies that comprehensively collect information on multiple types of risk factors. The International Lymphoma Epidemiology Consortium (InterLymph) may serve as a prototype for such studies.

Article Highlights

- Follicular lymphoma is the second most common subtype of NHL and the most common indolent one.
- Despite a significant amount of effort, the etiology and prognosis of FL remain poorly understood.
- A large number of potential risk factors have been suggested in the literature. However, there is a lack of consensus, and many suggested risk factors have not been rigorously validated in independent studies.
- More studies are needed to consolidate the existing findings and discover more FL risk factors. There is also need for research on personalized risk factors.

Declaration of Interest

The author declares no conflict of interest.

Bibliography

1. Howlader N, Noone AM, Krapcho M, et al. SEER Cancer Statistics Review, 1975-2008. National Cancer Institute. Bethesda, MD, based on November 2010 SEER data submission.
2. Ekstrom-Smedby K. Epidemiology and etiology of Non-Hodgkin lymphoma: a review. Acta Oncol. 2006; 45: 258-271.
3. Freedman A. Follicular lymphoma: 2011 update on diagnosis and management. Am J Hematol. 2011; 86: 768-775.
4. Solal-Celigny P, Cahu X, Cartron G. Follicular lymphoma prognostic factors in the modern era: what is clinically meaningful? Int J Hematol 2010; 92: 246-254.

5. Zhang Y, Dai Y, Zheng T, et al. Risk factors of Non-Hodgkin Lymphoma. Expert Opin Med Diagn. 2011; 5: 539-550.
6. Wang SS, Nieters A. Unraveling the interactions between environmental factors and genetic polymorphisms in non-Hodgkin lymphoma risk. Expert Rev Anticancer Ther. 2010; 10: 403-413.
7. Wang SS, Slager SL, Brennan P, et al. Family history of hematopoietic malignancies and risk of non-Hodgkin lymphoma (NHL): a pooled analysis of 10,211 cases and 11,905 controls from the International Lymphoma Epidemiology Consortium (InterLymph). 2007; 109: 3479-3488.
8. Hartge P, Wang SS, Bracci PM. et al. Non-Hodgkin Lymphoma. In Schottenfeld D, Fraumeni JFJ editors. Cancer Epidemiology and Prevention. Oxford University Press; New York: 2006. 898-918.
9. Zhang Y, Sanjose SD, Bracci PM. et al. Personal use of hair dye and the risk of certain subtypes of non-Hodgkin lymphoma. American Journal of Epidemiology. 2008; 167: 1321-1331.
10. Talamini R, Polesel J, Montella M, et al. Smoking and Non-Hodgkin's Lymphoma: Case-Control Study in Italy. International Journal of Cancer. 2005; 115: 606-610.
11. Morton LM, Holford TR, Leaderer B, et al. Cigarette smoking and risk of non-Hodgkin lymphoma subtypes among women. British Journal of Cancer. 2003; 89: 2087-2092.
12. Troy JD, Hartge P, Weissfeld JL, et al. Associations between anthropometry, cigarette smoking, alcohol consumption, and non-Hodgkin lymphoma in the

Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial. American Journal of Epidemiology. 2010; 171; 1270-1281.

13. Morton LM, Zheng T, Holford TR, et al. Alcohol consumption and risk of non-Hodgkin lymphoma: a pooled analysis. Lancet Oncol 2005; 6(7): 469-476.
14. Chang ET, Clarke CA, Canchola AJ, et al. Alcohol consumption over time and risk of lymphoid malignancies in the California Teachers Study cohort. American Journal of Epidemiology. 2010; 172: 1373-1383.
15. Pylypchuk RD, Schouten LJ, Goldbohm RA, et al. Body mass index, height, and risk of lymphatic malignancies: a prospective cohort study. American Journal of Epidemiology. 2009; 170: 297-307.
16. Cerhan JR, Janney CA, Vachon CM, et al. Anthropometric characteristics, physical activity, and risk of non-Hodgkin's lymphoma subtypes and B-cell chronic lymphocytic leukemia: a prospective study. Am J Epidemiol 2002;156:527–35.
17. Skibola CF, Holly EA, Forrest MS, et al. Body mass index, leptin and leptin receptor polymorphisms, and non-Hodgkin lymphoma. Cancer Epidemiol Biomarkers Prev 2004; 13:779–86
18. Cheng ET, Hjalgrim H, Smedby KE, et al. Body mass index and risk of malignant lymphoma in Scandinavian men and women. JNCI. 2005; 97: 210-8.
19. Lan Q, Zheng T, Chanock S, et al. Genetic variants in caspase genes and susceptibility to non-Hodgkin lymphoma. Carcinogenesis 2007; 28: 823-7.
20. Lan Q, Zheng T, Shen M, et al. Genetic polymorphisms in the oxidative stress pathway and susceptibility to non-Hodgkin lymphoma. Hum genet 2007; 121: 161-8.

21. Rothman N, Skibola CF, Wang SS, et al. Genetic variation in tnf and il10 and risk of non-Hodgkin lymphoma: a report from the Interlymph consortium. Lancet oncol 2006; 7: 27-38.
22. Bende RJ, Smit LA, van Noesel CJM. Molecular pathways in follicular lymphoma. Leukemia. 2007; 21: 18-29.
23. Dunphy CH. Gene expression profiling data in lymphoma and leukemia : review of the literature and extrapolation of pertinent clinical applications. Arch Pathol Lab Med. 2006; 130: 483-520.
24. Husson H, Carideo EG, Neuberg D, et al. Gene expression profiling of follicular lymphoma and normal germinal center B cells using cDNA arrays. Blood. 2002; 99: 282-289.
25. Conde L, Halperion E, Akers NK, et al. Genome-wide association study of follicular lymphoma identifies a risk locus at 6p21.32. Nature Genetics. 2010; 42: 661-664.
26. Smedby KE, Foo JN, Skibola CF, et al. GWAS of follicular lymphoma reveals allelic heterogeneity at 6p21.32 and suggests shared genetic susceptibility with diffuse large B-cell lymphoma. PLoS Genetics. 2011; 7: e1001378.
27. Hayslip J, Montero A. Tumor suppressor gene methylation in follicular lymphoma: a comprehensive review. Molecular Cancer. 2006; 5: 44.
28. McDonald HL, Gascoyne RD, Horsman D, et al. Involvement of the X chromosome in non-Hodgkin lymphoma. Genes Chromosomes Cancer. 2000; 28(3): 246-257.
29. Yang H, Chen CM, Yan P, et al. The androgen receptor gene is preferentially hypermethylated in follicular non-Hodgkin's lymphomas. Clin Cancer Res. 2003; 9(11): 4034-4042.

30. Koyama M, Oka T, Ouchida M, et al. Activated proliferation of B-cell lymphomas/leukemias with the SHP1 gene silencing by aberrant CpG methylation. Lab Invest 2003, 83(12):1849-1858.
31. Rossi D, Capello D, Gloghini A, et al. Aberrant promoter methylation of multiple genes throughout the clinico-pathologic spectrum of B-cell neoplasia. Haematologica 2004, 89(2):154-164.
32. Nakatsuka S, Takakuwa T, Tomita Y, et al. Hypermethylation of death-associated protein (DAP) kinase CpG island is frequent not only in B-cell but also in T- and natural killer (NK)/T-cell malignancies. Cancer Sci 2003, 94(1):87-91.
33. Stranks G, Height SE, Mitchell P, et al. Deletions and rearrangement of CDKN2 in lymphoid malignancy. Blood 1995, 85(4):893-901.
34. Pinyol M, Cobo F, Bea S, et al. p16(INK4a) gene inactivation by deletions, mutations, and hypermethylation is associated with transformed and aggressive variants of non-Hodgkin's lymphomas. Blood 1998, 91(8):2977-2984.
35. Baur AS, Shaw P, Burri N, et al. Frequent methylation silencing of p15(INK4b) (MTS2) and p16(INK4a) (MTS1) in B-cell and T-cell lymphomas. Blood 1999, 94(5):1773-1781.
36. Martinez-Delgado B, Robledo M, Arranz E, et al. Hypermethylation of p15/ink4b/MTS2 gene is differentially implicated among non-Hodgkin's lymphomas. Leukemia 1998, 12(6):937-941.
37. Li Y, Nagai H, Ohno T, et al. Aberrant DNA methylation of p57(KIP2) gene in the promoter region in lymphoid malignancies of B-cell phenotype. Blood 2002, 100(7): 2572-2577.

38. Solal-Celigny P, Roy P, Colombat P, et al. Follicular lymphoma international prognostic index. Blood. 2004; 104: 1258-1265.
39. Luminari S, Cox MC, Montanini A, et al. Prognostic tools in follicular lymphomas. Expert Rev Hematol. 2009; 2: 549-562.
40. Federico M, Bellei M, Marcheselli L, et al. Follicular Lymphoma International Prognostic Index 2: a new prognostic index for follicular lymphoma developed by the International Follicular Lymphoma Prognostic Factor Project. JNCI. 2009; 27: 4555-4562.
41. Arcaini L, Merli M, Passamonti F, et al. Validation of follicular lymphoma international prognostic index 2 (FLIPI2) score in an independent series of follicular lymphoma patients. Br J Haemotol. 2010; 149: 455-457.
42. Geyer SM, Morton LM, Habermann TM, et al. Smoking, alcohol use, obesity, and overall survival from non-Hodgkin lymphoma: a population-based study. Cancer. 2010; 116: 2993-3000.
43. Han X, Zheng T, Foss F, et al. Vegetable and fruit intake and non-Hodgkin lymphoma survival in Connecticut women. Leuk Lymphoma. 2010; 51: 1047-1054.
44. Thompson CA, Cerhan JR. Fruit and vegetable intake and survival from non-Hodgkin lymphoma: does an apple a day keep the doctor away? Leuk Lymphoma. 2010; 51: 963-964.
45. Dave SS, Wright G, Tan B, et al. Prediction of survival in follicular lymphoma based on molecular features of tumor0infiltrating immune cells. NEJM. 2004; 351: 2159-2169.

46. Tibshirani R. Immune signatures in follicular lymphoma. NEJM. 2005; 352: 1496-1497. Comment on NEJM 2004; 351: 2159-2169.
47. Glas AM, Kersten J, Delahaye LJML, et al. Gene expression profiling in follicular lymphoma to assess clinical aggressiveness and to guide the choice of treatment. Blood. 2005; 105: 301-307.
48. Lossos IS, Alizadeh AA, Diehn M, et al. Transformation of follicular lymphoma into diffuse large cell lymphoma: alternative patterns with increased or decreased expression of c-myc and its regulated genes. PNAS. 2002; 99: 8886-8891.
49. De Vos S, Hofmann WK, Grogan TM, et al. Gene expression profile of serial samples of transformed B-cell lymphomas. Lab Invest. 2003; 83: 271-285.
50. Elenitoba-Johnson KS, Jenson SD, Abbott RT, et al. Involvement of multiple signaling pathways in follicular lymphoma transformation: p38-mitogen-activated protein kinase as a target for therapy. PNAS 2003. 100: 7259-7264.
51. Janikova A, Tichy B, Supikova J, et al. Gene expression profiling in follicular lymphoma and its implication for clinical practice. Leukemia and Lymphoma. 2011; 52: 59-68.
52. Cerhan JR, Wang SS, Maurer MJ, et al. Prognostic significance of host immune gene polymorphisms in follicular lymphoma survival. Blood. 2007. 1009; 5439-5446.
53. Wrench D, Leighton P, Skibola CF, et al. SNP rs6457327 in the HLA region on chromosome 6p is predictive of the transformation of follicular lymphoma. Blood. 2011; 117: 3147-3150.
54. Han X, Li Y, Zhang Y, et al. Identification of predictive pathways for non-Hodgkin lymphoma prognosis. Cancer Informatics. 2010; 9: 281-292.

55. Ma S, Zhang Y, Huang J, et al. Identification of non-Hodgkin's lymphoma prognosis signatures using the CTGDR method. Bioinformatics. 2010; 26: 15-21.
56. Cheung KJJ, Johnson NA, Affleck JG, et al. Acquired TNFRSF14 mutations in follicular lymphoma are associated with worse prognosis. Cancer Research. 2010; 70: 9166-9174.
57. O'Shea D, O'Riain C, Gupta M, et al. Regions of acquired uniparental disomy at diagnosis of follicular lymphoma are associated with both overall survival and risk of transformation. Blood. 2009; 113: 2298-2301.
58. O'Riain C, O'Shea DM, Yang Y, et al. Array-based DNA methylation profiling in follicular lymphoma. Leukemia. 2009; 23: 1858-1866.
59. Ardeshna KM, Smith P, Norton A, et al. Long-term effect of a watch and wait policy versus immediate systemic treatment for asymptomatic advanced-stage non-Hodgkin lymphoma: a randomised controlled trial. Lancet. 2003; 362: 516-522.
60. Ardeshna KM, Smith P, Qian W, et al. An intergroup randomised trial of rituximab versus a watch and wait strategy in patients with stage II, III, IV, asymptomatic, non-bulky follicular lymphoma (grades 1, 2, and 3a). A preliminary analysis. Blood. 2010; 116: abstract 6.
61. Cheson BD. New agents in follicular lymphoma. Best Pract Res Clin Haematol. 2011; 24: 305-312.
62. Hitz F, Ketterer N, Lohri A, et al. Diagnosis and treatment of follicular lymphoma. Swiss Med Wkly. 2011; 141:w13247.
63. Johnson NA, Gascoyne RD. Gene expression signatures in follicular lymphoma: are they ready for the clinic? Haematologica. 2008; 93: 982-987.