

The prognosis of follicular lymphomas: the F2-project

M. FEDERICO¹
M. BELLEI¹
B. PRO²
A. LOPEZ-GUILLERMO³
U. VITOLO⁴
S. LUMINARI¹
S. PILERI⁵
P. SOLAL-CÉLIGNY⁶

ON BEHALF OF THE IFLPFP

¹Univ. Modena e Reggio Emilia, Modena, Italy;

²Univ. Texas, MD Anderson Cancer Center, Houston, USA;

³Hosp. Clin., Barcelona, Spain;

⁴Osp. S. Giovanni B., Torino, Italy

⁵Univ. Bologna, Bologna, Italy;

⁶Centre J. Bernard, Le Mans, France

Follicular lymphoma (FL) accounts for 10%–15% of non-Hodgkin's lymphomas in western countries.¹

Although patients with FL experience a relatively indolent course and usually exhibit dramatic responses to initial therapy, they should be considered affected by a fatal malignancy. There is a tendency to relapse over time, the response to salvage treatments is of shorter duration after every relapse, and patients eventually die of disease-related causes.² Several treatment approaches are offered to patients with FL;^{2–4} however, criteria to rationalize treatment decisions are still lacking in many instances.

So far, a variety of studies involving patients with FL have targeted the evaluation of prognosis. Different demographic, clinical, and biological factors have shown a prognostic role, including age, sex, stage, tumor burden, bone marrow involvement, systemic symptoms, performance status, serum lactate dehydrogenase (LDH) level, anemia, erythrocyte sedimentation rate (ESR), and β -2 microglobulin. More recently, with the advent of microarray technology, promising data are also available about the prognostic role of lymphoma cell genomic expression and the role of the microenvironment immune signature.

Attempts to define prognosis in FL began in the late 1970s^{5–7} when Leonard *et al.*⁶ first constructed a prognostic index based on age, sex, stage, hemoglobin level, and performance status, which separates patients into groups at high, intermediate, and low risk. Romaguera *et al.*⁵ have in turn developed an index based on sex and on tumor burden defined by using the number of extranodal sites, the size of involved lymph nodes, and the degree of bone marrow infiltrate.

Then, when the International Prognostic Index (IPI) was defined in 1993 for aggressive lymphomas⁸, it was also applied to low-grade lymphomas, leading to conflicting results^{9,10}, and the need for a prognostic index specifically designed for FL emerged.

In response to this need, a large study

based on 987 patients was performed by the Italian Lymphoma Intergroup (ILI).¹¹

The ILI index (Table 1) was based on age, gender, systemic symptoms, number of extranodal sites, ESR, and LDH. These six variables allowed the definition of a prognostic model with three risk groups associated with different 5- and 10-year survival rates. Patients with 0 or 1 risk factors were considered to be at low risk; those with 2 unfavorable variables were at intermediate risk; and those with 3 or more unfavorable variables were at high risk.

More recently, the International Follicular Lymphoma Prognostic Factor Project (IFLPFP) delineated the definition of the FLIPI index (Table 1), developed as a result of a large international cooperative effort.¹² The score was defined on a training series of 1,795 patients and was based on five variables: age, Ann Arbor stage, hemoglobin level, number of nodal site areas, and serum LDH. Based on the final model, patients with 0 or 1 risk factors were characterized by a 5- and 10-year overall survival (OS) of 91% and 71%, respectively; patients with 2 risk factors had an intermediate 5- and 10-year OS of 78% and 51%, respectively; and those with 3 or more risk factors, which represented 27% of all cases, had the worst 5- and 10-year OS of 53% and 36%, respectively.

IPI, ILI, and FLIPI models have been recently compared in a large group of patients to determine the relative merits of each approach.¹³ Overall concordance of the three systems was 54%; concordance was 37% for low-risk groups, 10% for intermediate groups, and 36% for high-risk groups. All three prognostic scores are easily applicable in clinical practice because they include variables that are easy to calculate. In the comparison study, the FLIPI score classified more patients in the high-risk group than did the IPI and ILI, even when only younger patients were considered. However, the high-risk group identified by the ILI system, although numerically less consistent, was characterized by a worse

Table 1. Comparison between ILI and FLIPI indexes and distribution of patients according to risk, based on originally published data.

ILI ¹¹			FLIPI ¹²	
		Model definition		
Age >60 years			Age >60 years	
ENS > 2			Elevated LDH	
Elevated LDH			Stage III-IV	
-			Nodal sites > 4	
-			-	
Male gender			Hb level < 12 g/dl	
-			-	
B-symptoms			-	
ESR > 30 mm/h			-	
		Period of diagnosis		
1985-1996			1985-1992	
		Initial study population		
987 pts			5120 pts	
		Final study sample		
429 pts			1795 pts	
		Median follow-up		
54 m			90m	
		Patients' distribution		
Low risk (0-1)	64%		Low risk (0-1)	36%
Intermediate risk (2)	23%		Intermediate risk (2)	37%
High risk (3-5)	13%		High risk (3-5)	27%
		5 and 10 year OS according to risk		
Low risk (0-1)	90% - 65%		Low risk (0-1)	91%-71%
Intermediate risk (2)	75% - 54%		Intermediate risk (2)	78%-51%
High risk (3-5)	38% - 11%		High risk (3-5)	53%-36%

ENS, extranodal sites, LDH, lactic dehydrogenase; Hb, hemoglobin; ESR, erythrocyte sedimentation rate.

prognosis than the corresponding IPI and FLIPI high-risk groups. This finding confirms that ILI may have a more relevant role than other prognostic models in selecting patients with poor prognosis.

Notwithstanding the huge number of patients considered in these studies, all mentioned prognostic scores (IPI, ILI, and FLIPI) are based on the retrospective collection of archive data. This approach can introduce biases that can hamper final results. A first problem is the selection of patients, which can be influenced by specific institution policy and patient- or physician-related factors. Furthermore, some promising prognostic factors, such as β 2-microglobulin or ESR, were not included in the published indexes because they were available only in a small number of patients. Finally, the results of a retrospective analysis targeting survival depend also on the type of treatment; with the recent advent of new drugs, such as monoclonal antibodies and purine analogs that can be used in elderly patients, the role of some established prognostic factors may have changed.

For these reasons, the F2-study was launched at the beginning of 2003. The F2-study was conceived as a complement to the IFLFPF study, and its purpose was to validate the FLIP Index and to verify whether a prognostic collection of data would allow the development of a more accurate prognostic index. The study was

designed as a prospective collection of information potentially useful for predicting the prognosis of newly diagnosed patients with FL. Study proponents reasoned that a prospective registration of patients in a short period of time would allow collection of an exhaustive set of clinical data and biological information, thus limiting selection biases as much as possible.

Patients and Methods

Patients were registered in the study regardless of planned treatment, including a watch and wait policy. Registration was performed on-line on a key-restricted, accessible web database. Patients eligible for the F2-study satisfied the following inclusion criteria: newly diagnosed FL; histologically confirmed FL diagnosis according to REAL/WHO classification (any grade); and age over 18. No exclusion criteria were set. Inclusion criteria were checked at the Trial Office (University of Modena and Reggio Emilia, Modena).

Statistical considerations

If the risk factor has a 10% prevalence, the 5-year survival of the remaining subjects is 70%, and the odds ratio is 2 for death with the risk factor compared to that without, then there will be 80% power to detect a statistically significant effect of the risk factor on survival with a sample size of 750 patients. Because of

Table 2. Characteristics of 941 patients with complete data.

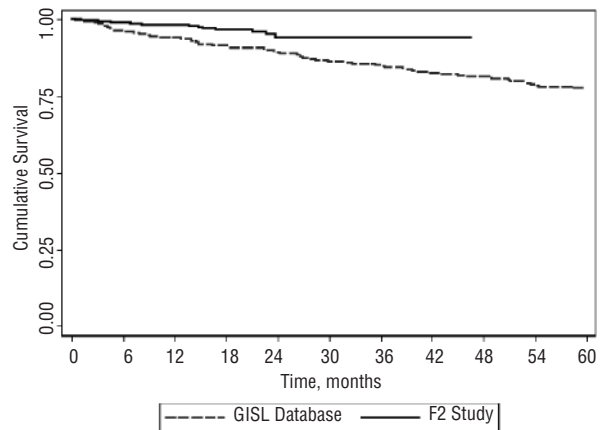
	<i>n</i>	%
Age		
<60 years	546	58
≥60 years	345	42
Median age		
Gender		
Male	473	50
Female	468	50
Systemic symptoms presence	107	11
ECOG-PS		
0	702	75
1	202	21
2+	37	4
Stage		
I	168	18
II	125	13
III	220	23
IV	428	46
Type of disease		
Nodal (± spleen ± BM)	763	81
Nodal/Extranodal	123	13
Pure extranodal	55	6
FLIPI		
Low	447	47
Intermediate	288	31
High	206	22

the interrelationship between risk factors, a total sample size of 900 patients is likely to yield similar power in a multivariate analysis, which allows for the effect of several risk factors. Thus, considering an approximate 10% drop-out rate of patients due to ineligibility for any cause and patients lost during the 5 years of follow-up, a total accrual of 1,000 cases was planned.

The primary endpoint of the study was 5-year OS, and the second endpoint was 5-year event-free survival. Additional endpoints were remission rate with initial therapy, 5-year progression-free survival, and 5-year treatment-free survival.

Results

Between January 2003 and May 2005, 69 European and U.S. institutions contributed to the F2-study, for a total of 1,093 patients. After registration, 19 cases were excluded for revised diagnosis, violation of inclusion criteria, or data unavailability. At present, 941 (90%) of the 1,074 patients eligible for the study have all the information requested for previously developed prognostic indexes assessment (age, gender, presence of systemic symptoms, stage, number of nodal and extranodal sites of disease, ECOG-PS, LDH, and ESR levels, and hemoglobin level); moreover, information on β -2 microglobulin levels and bcl-2 status were

**Figure 1. Overall survival of F2-study patients and of GISL-enrolled patients.**

available in 91% (860 patients) and 65% (615 patients) of cases, respectively.

In addition to consisting of an extremely broad population with a high level of completeness, the F2-study is also characterized by a homogeneous set of patients diagnosed and staged according to current guidelines. A summary of patient characteristics is reported in Table 2.

After a median follow-up of 9 months, the 2-year OS is 95%. Compared with survival observed in 374 patients with FL enrolled in different prospective trials by the Gruppo Italiano per lo Studio dei Linfomi (G.I.S.L.) between 1990 and 2004 (*unpublished data*), the outcome of patients enrolled in the F2-study seems more promising (Figure 1).

Conclusions

In conclusion, the F2-study allowed the collection of an exhaustive set of high-quality clinical and biological data in a very short period of time and would permit development of a prognostic index that would be useful for clinicians as a tool to deliver the most suitable treatment with the most successful result for the patient.

Acknowledgments

The authors wish to thank the *Fondazione Cassa di Risparmio di Modena* and the *Associazione Angela Serra per la ricerca sul cancro* for supporting the study

References

1. A clinical evaluation of the International Lymphoma Study Group classification of non-Hodgkin's lymphoma. The Non-Hodgkin's Lymphoma Classification Project. *Blood*, 1997;89:3909-18.
2. Horning SJ, Natural history of and therapy for the indolent non-Hodgkin's lymphomas. *Semin Oncol*, 1993. 20(5 Suppl 5): p. 75-88.
3. Ardeschna KM, 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: p. 516-22.
4. Brandt, L., et al., A systematic overview of chemotherapy effects in indolent non-Hodgkin's lymphoma. *Acta Oncol* 2001;40: 213-23.
 5. Romaguera, J.E., et al., Multivariate analysis of prognostic factors in stage IV follicular low-grade lymphoma: a risk model. *J Clin Oncol* 1991;9:762-9.
 6. Leonard RC, et al., The identification of discrete prognostic groups in low grade non-Hodgkin's lymphoma. The Scotland and Newcastle Lymphoma Group Therapy Working Party. *Ann Oncol* 1991;2: 655-62.
 7. Decaudin D., et al., Low-grade stage III-IV follicular lymphoma: multivariate analysis of prognostic factors in 484 patients--a study of the groupe d'Etude des lymphomes de l'Adulte. *J Clin Oncol* 1999;17:2499-505.
 8. A predictive model for aggressive non-Hodgkin's lymphoma. The International Non-Hodgkin's Lymphoma Prognostic Factors Project. *N Engl J Med* 1993;329:987-94.
 9. Cameron DA, et al., Identification of prognostic groups in follicular lymphoma. The Scotland and Newcastle Lymphoma Group Therapy Working Party. *Leuk Lymphoma* 1993;10:89-99.
 10. Lopez-Guillermo A, et al. Applicability of the International Index for aggressive lymphomas to patients with low-grade lymphoma. *J Clin Oncol* 1994;12:1343-8.
 11. Federico M, et al, Prognosis of follicular lymphoma: a predictive model based on a retrospective analysis of 987 cases. *Intergruppo Italiano Linfomi. Blood* 2000;95:783-9.
 12. Solal-Celigny P, et al, Follicular lymphoma international prognostic index. *Blood* 2004;104:1258-65.
 13. Perea G, et al. Prognostic indexes in follicular lymphoma: a comparison of different prognostic systems. *Ann Oncol* 2005;16: 1508-13.