A PREDICTIVE MODEL FOR AGGRESSIVE NON-HODGKIN'S LYMPHOMA

THE INTERNATIONAL NON-HODGKIN'S LYMPHOMA PROGNOSTIC FACTORS PROJECT*

Abstract *Background.* Although many patients with intermediate-grade or high-grade (aggressive) non-Hodgkin's lymphoma are cured by combination chemotherapy, the remainder are not cured and ultimately die of their disease. The Ann Arbor classification, used to determine the stage of this disease, does not consistently distinguish between patients with different long-term prognoses. This project was undertaken to develop a model for predicting outcome in patients with aggressive non-Hodgkin's lymphoma on the basis of the patients' clinical characteristics before treatment.

Methods. Adults with aggressive non-Hodgkin's lymphoma from 16 institutions and cooperative groups in the United States, Europe, and Canada who were treated between 1982 and 1987 with combination-chemotherapy regimens containing doxorubicin were evaluated for clinical features predictive of overall survival and relapse-free survival. Features that remained independently significant in step-down regression analyses of survival were incorporated into models that identified groups of patients of all ages and groups of patients no more than 60 years old with different risks of death.

Results. In 2031 patients of all ages, our model, based on age, tumor stage, serum lactate dehydrogenase concentration, performance status, and number of extranodal disease sites, identified four risk groups with predicted five-year survival rates of 73 percent, 51 percent, 43 percent, and 26 percent. In 1274 patients 60 or younger, an age-adjusted model based on tumor stage, lactate dehydrogenase level, and performance status identified four risk groups with predicted five-year survival rates of 83 percent, 69 percent, 46 percent, and 32 percent. In both models, the increased risk of death was due to both a lower rate of complete responses and a higher rate of relapse from complete response. These two indexes, called the international index and the age-adjusted international index, were significantly more accurate than the Ann Arbor classification in predicting long-term survival.

Conclusions. The international index and the age-adjusted international index should be used in the design of future therapeutic trials in patients with aggressive non-Hodgkin's lymphoma and in the selection of appropriate therapeutic approaches for individual patients. (N Engl J Med 1993;329:987-94.)

COMBINATION chemotherapy has transformed aggressive non-Hodgkin's lymphoma from a fatal disease into one that is often curable. However, many patients still die of their disease, underscoring the need for more accurate methods of prospectively identifying patients with different long-term prognoses. The identification of those at "high" or "low" risk could have important therapeutic implications. Patients at high risk who are not effectively treated with current regimens may benefit from new experimental approaches, whereas those at low risk may do well with standard therapy but sustain severe toxic reactions without additional benefit if they are treated with experimental regimens. The identification of different risk groups would also aid in the design and interpretation of therapeutic trials.

The tumor stage of patients with aggressive non-Hodgkin's lymphoma is currently determined with the Ann Arbor classification, which was originally developed for Hodgkin's disease.[1] This classification emphasizes the distribution of nodal disease sites because Hodgkin's disease commonly spreads through contiguous groups of lymph nodes.[1] Since the patterns of disease spread in Hodgkin's disease and non-Hodgkin's lymphoma are different, it is not surprising that the Ann Arbor classification system is less accurate in identifying prognostic subgroups of patients with aggressive non-Hodgkin's lymphoma.[2]

In previous analyses of relatively small numbers of patients with this disease, a variety of clinical characteristics were consistently associated with outcome: the age at diagnosis, the presence or absence of systemic (B) symptoms, performance status, the serum lactate dehydrogenase (LDH) concentration, the number of nodal and extranodal sites of disease, tumor size, and the distinction between localized disease (Ann Arbor stage I or II) and advanced disease (stage III or IV).[3-13] These features were thought to reflect the tumor's growth and invasive potential (LDH level, tumor stage, tumor size, number of nodal and extranodal sites of disease, and the presence or absence of bone marrow involvement), the patient's response to the tumor (performance status and status for B symptoms), and the patient's ability to tolerate intensive therapy (performance status, bone marrow involvement, and age). Many investigators identified a subgroup of clinical features that remained independently

*The following persons are members of the International Non-Hodgkin's Lymphoma Prognostic Factors Project: M.A. Shipp (Dana–Farber Cancer Institute and Harvard Medical School, Boston) and D.P. Harrington (Dana–Farber Cancer Institute and Harvard Medical School, Boston; Eastern Cooperative Oncology Group, Denver), chairpersons; J.R. Anderson (Nebraska Lymphoma Study Group, Omaha; Cancer and Leukemia Group B, Lebanon, N.H.); J.O. Armitage (Nebraska Lymphoma Study Group, Omaha); G. Bonadonna (Istituto Nazionale Tumori, Milan, Italy); G. Brittinger (Bundesministerium für Forschung und Technologie Study, Essen, Germany); F. Cabanillas (M.D. Anderson Cancer Center, Houston); G.P. Canellos (Dana–Farber Cancer Institute and Harvard Medical School, Boston); B. Coiffier (Groupe d'Etude des Lymphomes de l'Adulte, Paris); J.M. Connors (British Columbia Cancer Agency, Vancouver); R.A. Cowan and D. Crowther (Manchester Lymphoma Group, Manchester, United Kingdom); S. Dahlberg (Southwest Oncology Group, San Antonio, Tex.); M. Engelhard (Bundesministerium für Forschung und Technologie Study, Essen, Germany); R.I. Fisher (Southwest Oncology Group, San Antonio, Tex.); C. Gisselbrecht (Groupe d'Etude des Lymphomes de l'Adulte, Paris); S.J. Horning (Stanford University, Stanford, Calif.); E. Lepage (Groupe d'Etude des Lymphomes de l'Adulte, Paris); T.A. Lister (St. Bartholomew's Hospital, London); J.H. Meerwaldt (Lymphoma Cooperative Group, European Organisation for Research on Treatment of Cancer, Brussels, Belgium); E. Montserrat (Hospital Clinic i Provincial de Barcelona, Barcelona, Spain); N.I. Nissen (Finsen Institute, Copenhagen, Denmark); M.M. Oken (Eastern Cooperative Oncology Group, Denver); B.A. Peterson (Cancer and Leukemia Group B, Lebanon, N.H.); C. Tondini (Istituto Nazionale Tumori, Milan, Italy); W.A. Velasquez (M.D. Anderson Cancer Center, Houston); and B.Y. Yeap (Dana–Farber Cancer Institute and Harvard Medical School, Boston; Eastern Cooperative Oncology Group, Denver). Address reprint requests to Dr. Margaret A. Shipp at the Dana–Farber Cancer Institute, 44 Binney St., Boston, MA 02115.

Supported in part by the Arntraud C. Arlinger Fund.

significant in multivariate analyses of their patients and used this subgroup to develop prognostic models that would predict a given patient's risk of death.[4-11,13] Although the specific clinical features in these models differed, all models included the measurements of disease volume and extent of tumor involvement at presentation. To develop a better prognostic-factor model for aggressive non-Hodgkin's lymphoma, 16 institutions and cooperative groups in the United States, Europe, and Canada participated in the project described below.

Methods

Characteristics of the Patients

Participating centers submitted data on each eligible patient included in electronic files.

Adult patients were eligible for this study if they had diffuse mixed, diffuse large-cell, or large-cell immunoblastic lymphoma (International Working Formulation[14] categories F, G, and H); diffuse centroblastic–centrocytic, centroblastic, immunoblastic, or unclassified high-grade lymphoma (Kiel classification[15]); or diffuse mixed lymphocytic–histiocytic or diffuse histiocytic lymphoma (Rappaport classification[16]). All the patients were treated with a combination-chemotherapy regimen containing doxorubicin as part of a phase 2 or 3 study between 1982 and 1987. The inclusion of only patients who had completed therapy by 1987 ensured a minimum of 3 years of follow-up for all patients and a median of 4½ years of follow-up for surviving patients. The stages of the tumors were determined and their pathologic characteristics were reviewed according to guidelines at the participating institutions.

The clinical features evaluated for potential prognostic importance were sex, age, tumor stage, performance status, B symptoms, sites of lymphomatous involvement, number of extranodal disease sites, size of the largest tumor, and serum concentrations of LDH, albumin, and β_2-microglobulin. The Ann Arbor stage of the tumor was designated as I, II non-bulky (largest tumor dimension, <10 cm), II bulky (largest dimension, ≥10 cm), III, or IV. Performance status was assessed according to the Eastern Cooperative Oncology Group scale, in which 0 indicated that the patient had no symptoms; 1, the patient had symptoms but was ambulatory; 2, the patient was bedridden less than half the day; 3, the patient was bedridden half the day or longer; and 4, the patient was chronically bedridden and required assistance with the activities of daily living. Performance status was classified as 0 or 1 (the patient was ambulatory) or 2, 3, or 4 (the patient was not ambulatory) (equivalent Karnofsky scores, ≥80 and ≤70). B symptoms were defined as recurrent fever (temperature, >38.3°C [101°F]), night sweats, or the loss of more than 10 percent of body weight. The recorded sites of extranodal lymphomatous involvement included the bone marrow, gastrointestinal tract, liver, lung, central nervous system, and other sites; the numbers of extranodal disease sites were recorded as 0, 1, or more than 1. Splenic involvement was also recorded. The largest dimension of the largest site of bulky disease was measured and reported as being less than 10 cm or as 10 cm or more. The serum LDH level was expressed as the ratio of the measured value to the upper limit of the normal range reported in the laboratory of each participating institution.

Assessment of Response

A complete response to treatment was defined by the participating institutions as the disappearance of all clinical evidence of disease and the normalization of all laboratory values and radiographic results that had been abnormal before treatment. The relapse-free survival of patients with complete responses was measured as the interval between the end of treatment and relapse or death or the date of the last follow-up evaluation in patients who had no relapse. Survival was measured as the interval between the beginning of treatment and death or the date of the last follow-up evaluation.

Statistical Analysis

The univariate associations between response and individual clinical features were analyzed with Fisher's exact test for two-by-k tables.[17] Relapse-free survival among patients with complete responses and overall survival among all patients were estimated with the method of Kaplan and Meier.[18] The univariate associations between individual clinical features and overall survival and relapse-free survival were determined with the log-rank test.[19]

Since all centers had not originally gathered all the requested information on their patients, data on several prognostic factors were missing from the patients' files. No data on outcome were missing. Missing data were dealt with by carrying out "complete case" analyses, in which patients were excluded from particular analyses if their files did not contain data on the required variables. This method did not bias analysis, since the availability of data at each center was determined by the data collection at the time of treatment rather than by the eventual outcome.

Features independently associated with overall survival and relapse-free survival were identified in multivariate analyses by proportional-hazards regression.[20] Step-down regression methods were used to build parsimonious statistical models for the association of prognostic factors with overall survival and relapse-free survival among patients with complete responses. Time-dependent death rates (hazard functions) were estimated according to the nonparametric kernel smoothing methods of Gray.[21]

Results

Univariate Analysis of Predictive Features

The presenting characteristics of the 3273 patients with tumors in Ann Arbor stages I through IV who were included in the analysis are shown in Table 1. Of these patients, 66 percent had complete responses. The five-year relapse-free survival rate among the patients with complete responses was 59 percent, and the five-year overall survival among all patients was 52 percent (Table 2). The associations between the patients' characteristics and the response rate, relapse-free survival among those with complete responses, and overall survival are shown in Table 2. Given the size of the study population, it is not surprising that the majority of the listed clinical characteristics were significantly associated with outcome.

Independent Prognostic Factors and the Prognostic-Factor Model

Information on seven prognostic factors (age at diagnosis, performance status, serum LDH level, Ann Arbor stage, tumor size, number of extranodal disease sites, and presence or absence of B symptoms) that had been associated with outcome in many previous studies[3-13] was complete for 1872 patients, of whom 1385 (74 percent) were randomly selected as a training sample in which to identify independent prognostic factors to form a model. A training sample of nearly 75 percent was chosen because it would be large enough to detect a 20 percent increase in the relative risk of death associated with even a relatively rare characteristic.

The step-down regression analysis of overall survival in the training sample evaluated 12 variables: the 7 prognostic factors mentioned above, 4 individual sites of extranodal disease (bone marrow, liver, lung, and central nervous system), and the spleen. These 12 variables included all those in our data set except sex (which was not associated with survival), gastrointestinal involvement (which had a statistically significant but clinically unimportant association with survival [Table 2]), and serum β_2-microglobulin and albumin levels (on which we had insufficient data). In the regression analysis, age was coded as 60 years or less or more than 60 years because this dichotomy was most commonly used in previous analyses and because patients 60 or younger were the most likely candidates for intensive experimental therapy. The Ann Arbor stage and the number of extranodal disease sites were coded so that the individual categories and all natural dichotomous groupings for each variable were included (e.g., Ann Arbor stage I vs. II vs. III vs. IV; stage I vs. stages II through IV; stage I or II nonbulky disease vs. stage II bulky disease, stage III, or stage IV; and stage I or II vs. stage III or IV.) The most discriminating cutoff point for serum LDH was determined by applying classification and regression trees (which separate patients into homogeneous subgroups)[22] to martingale residuals (the differences between the number of events observed and the number predicted by the model).[23] The most predictive cutoff point for LDH was a level 1.2 times normal (≤1.2 times normal vs. >1.2 times normal) in both univariate and multivariate analyses; however, a level 1 times normal (≤1 vs. >1) was chosen as the cutoff point because it was almost as predictive as 1.2 times normal and easier to use.

Table 1. Characteristics of 3273 Patients Presenting with Aggressive Non-Hodgkin's Lymphoma.

CHARACTERISTIC	No. (%)*
Sex	
Male	2049 (63)
Female	1224 (37)
Age†	
≤60 yr	1931 (59)
>60 yr	1340 (41)
Unknown	2
Ann Arbor stage	
I	246 (8)
II, tumor <10 cm	492 (15)
II, tumor ≥10 cm	261 (8)
II, tumor size unknown	120 (4)
III	690 (21)
IV	1462 (45)
Unknown	2
Site of lymphomatous involvement	
Bone marrow	712 (22)
Gastrointestinal tract	528 (16)
Liver	350 (11)
Lung	290 (9)
Central nervous system	59 (2)
Other extranodal site	1210 (37)
Spleen	426 (13)
Extranodal involvement	
None	995 (30)
1 site	1241 (38)
>1 site	974 (30)
Unknown	63
Dimension of largest tumor‡	
<10 cm	1863 (57)
≥10 cm	777 (24)
Unknown	633
Performance status§	
0: Fully active	873 (27)
1: Ambulatory	1128 (34)
2: Bedridden <50% time	409 (12)
3: Bedridden ≥50% time	140 (4)
4: Completely bedridden	72 (2)
Unknown	651
B symptoms	
Absent	1941 (59)
Present	1326 (41)
Unknown	6
Serum LDH level¶	
≤1× normal	1208 (37)
>1× normal	1320 (40)
Unknown	745
Serum albumin level‖	
≥1× normal	652 (20)
<1× normal	269 (8)
Unknown	2352
Serum β_2-microglobulin level**	
≤1× normal	110 (3)
>1× normal	105 (3)
Unknown	3058

*Because of rounding, not all percentages total 100.

†Median, 56 years; range, 16 to 92.

‡Median, 7 cm; range, 1 to 34.

§The scale for performance status is that of the Eastern Cooperative Oncology Group, described in the Methods section.

¶Median, 1.0 times normal; range, 0.2 to 30.3.

‖Median, 1.1 times normal; range, 0.5 to 1.8.

**Median, 1.0 times normal; range, 0.3 to 8.0.

The five pretreatment characteristics that remained independently significant in the analysis of the training sample were age (≤60 vs. >60 years), tumor stage (stage I or II [localized disease] vs. stage III or IV [advanced disease]), the number of extranodal sites of disease (≤1 vs. >1), performance status (0 or 1 vs. ≥2), and serum LDH level (≤1 times normal vs. >1 times normal) (Table 3). These five features were used to design a model to predict an individual patient's risk of death — the international index. Since the relative risks associated with each of the independently significant risk factors were comparable (Table 3), the relative risk of death could be characterized by summing the number of risk factors present at diagnosis. Risk groups were defined by comparing the relative risk of death in patients with each possible number of presenting risk factors (0, 1, 2, 3, 4, or 5) and combining categories with similar relative risks (e.g., 0 with 1 or 4 with 5). Patients were then assigned to one of four risk groups on the basis of their number of presenting risk factors: 0 or 1, low risk; 2, low intermediate risk; 3, high intermediate risk; or 4 or 5, high risk. The survival curves and death rates over time for the four risk groups in the training sample are shown in Figure 1. To provide a basis for comparison, the survival curve and death rate for all 3273 patients in the study are also included (Fig. 1).

The prognostic-factor model was then applied to

a validation sample of patients, which contained the remaining patients with complete data on the seven specified variables (487 of the 1872 patients with complete data) and the other patients with complete data on only the five variables in the final model (159 additional patients; total, 646). The model was equally predictive in the validation sample, identifying four groups of patients at low, low intermediate, high intermediate, or high risk of death (Fig. 1). In the training and validation samples, the risk of death was increased primarily in the first three to four years after diagnosis (Fig. 1, right panels).

Table 2. Outcome According to the Patients' Characteristics.*

Characteristic	Complete Response		Relapse-Free Survival		Overall Survival	
	Rate (%)	P value	5-yr rate (%)	P value	5-yr rate (%)	P value
All patients	66		59		52	
Sex						
Male	66	0.849	60	0.112	52	0.455
Female	66		57		53	
Age						
≤60 yr	68	<0.001	67	<0.001	60	<0.001
>60 yr	62		49		41	
Ann Arbor stage						
I	94	<0.001	75	<0.001	79	<0.001
II, tumor <10 cm	85		68		68	
II, tumor ≥10 cm	77		69		64	
II, tumor size unknown	65		69		62	
III	67		56		52	
IV	52		48		40	
I or II	83	<0.001	70	<0.001	69	<0.001
III or IV	57		51		44	
Site of involvement						
Bone marrow						
Absent	69	<0.001	62	<0.001	56	<0.001
Present	55		40		37	
Gastrointestinal tract						
Absent	69	0.004	59	0.023	54	0.009
Present	62		68		51	
Liver						
Absent	68	<0.001	60	<0.001	55	<0.001
Present	49		46		34	
Lung						
Absent	68	<0.001	60	0.022	54	<0.001
Present	50		46		39	
Central nervous system						
Absent	67	0.001	59	0.046	53	<0.001
Present	46		20		29	
Spleen						
Absent	67	<0.001	62	<0.001	55	<0.001
Present	54		39		40	
Extranodal site						
None	75	<0.001	61	<0.001	61	<0.001
1 site	69		62		56	
>1 site	53		52		40	
≤1 site	72	<0.001	61	<0.001	58	<0.001
>1 site	53		52		40	
Dimension of largest tumor						
<10 cm	72	<0.001	58	0.005	55	0.005
≥10 cm	60		68		51	
Performance status						
Ambulatory (0 or 1)	70	<0.001	58	0.019	55	<0.001
Not ambulatory (2–4)	45		54		35	
B symptoms						
Absent	75	<0.001	62	<0.001	61	<0.001
Present	52		52		40	
Serum LDH level						
≤1× normal	81	<0.001	65	<0.001	67	<0.001
>1× normal	60		53		44	
Serum albumin level						
≥1× normal	76	<0.001	64	0.393	60	<0.001
<1× normal	55		51		37	
Serum β_2-microglobulin level						
≤1× normal	75	0.004	59	0.739	65	<0.001
>1× normal	55		75		39	

*Additional data on the number of patients treated at each institution, treatment regimens, age, stage criteria, and complete-response and five-year survival rates are available from the authors.

Since the training and validation samples had comparable outcomes, we combined these two groups into a single group for further detailed analysis (Fig. 1, bottom panels; Table 4, international index, all patients). The four risk groups had distinctly different rates of complete response, relapse-free survival, and overall survival (Table 4). For example, the low-risk group had a complete-response rate of 87 percent and a five-year overall survival of 73 percent, whereas the high-risk group had a complete-response rate of only 44 percent and a five-year overall survival of only 26 percent (Table 4, international index, all patients).

The Significance of Age in Prognosis

Since the two age groups (≤60 vs. >60 years) had significantly different outcomes (Tables 2 and 3) and the age limit for patients treated by most intensive experimental regimens for non-Hodgkin's lymphoma is 60 years, we also developed an age-adjusted model for younger patients — the age-adjusted international index. Three of the previously identified risk factors — tumor stage, performance status, and LDH level — remained independently significant prognostic factors among the patients in the training sample who were 60 or younger (885 patients) (Table 3). Since the relative risks of death associated with the three risk factors were comparable (Table 3), a younger patient could also be assigned to a risk group by counting the number of risk factors present at diagnosis. The age-adjusted international index was similarly predictive in the training and vali-

Table 3. Factors Independently Prognostic of Overall Survival in the Training Sample.

Factor	Relative Risk	P Value
All patients (n = 1385)		
Age (≤60 vs. >60)	1.96	<0.001
Serum LDH (≤1× normal vs. >1× normal)	1.85	<0.001
Performance status (0 or 1 vs. 2–4)	1.80	<0.001
Stage (I or II vs. III or IV)	1.47	<0.001
Extranodal involvement (≤1 site vs. >1 site)*	1.48	<0.001
Patients ≤60 years old (n = 885)		
Stage (I or II vs. III or IV)	2.17	<0.001
Serum LDH (≤1× normal vs. >1× normal)	1.95	<0.001
Performance status (0 or 1 vs. 2–4)	1.81	<0.001

*This was the only factor that did not retain independent prognostic significance in patients ≤60 years old (≤1 site vs. >1 site: relative risk, 1.20; P = 0.134).

dation samples (in 885 and 389 patients, respectively, or a total of 1274), justifying our combining these two groups for further analysis (Fig. 2; Table 4, age-adjusted index, patients ≤60). The younger patients (≤60 years) were assigned to four risk groups according to the number of risk factors at presentation (0, 1, 2, or 3) (Fig. 2 and Table 4). The younger patients in the low-risk group had a complete-response rate of 92 percent and a five-year overall survival of 83 percent, whereas those in the high-risk group had a complete-response rate of only 46 percent and a five-year overall survival of only 32 percent (Table 4). As was true of all patients (Table 4, international index, all patients), the increased risk of death among the younger patients was due to both a lower complete-response rate and lower relapse-free survival among those with complete responses.

To define the differences between the younger and older patients more specifically, the two age groups were compared by the age-adjusted international index (Table 4). The distribution of the younger patients among the four risk groups was similar to the distribution of the older patients among these groups (Table 4). Although the older patients (>60 years) had complete-response rates that were similar to or only slightly lower than those of the younger patients, the older patients with complete responses had much lower rates of relapse-free survival (Table 4). These data suggested that increased numbers of older patients died of lymphoma rather than of unrelated causes. Consistent with this observation was the finding that the older patients also had an observed death rate that was substantially higher than that of an age-matched cohort (data not shown).[24]

Relapse-free Survival among Patients with Complete Responses

If patients with an increased risk of relapse from complete response could be identified before relapse, they might be candidates for experimental approaches to consolidation therapy, such as high-dose chemoradiotherapy with infusion of peripheral-blood stem cells or bone marrow support. For this reason, we also identified the presenting clinical features most closely associated with the risk of relapse from complete response: age (≤60 vs. >60 years: relative risk, 1.80; P<0.001), tumor stage (I or II vs. III or IV: relative risk, 1.79; P<0.001), and serum LDH level (≤1 times normal vs. >1 times normal: relative risk, 1.47; P<0.001).

International Index and Ann Arbor Stage

Since features other than the Ann Arbor stage were independently associated with overall survival in our

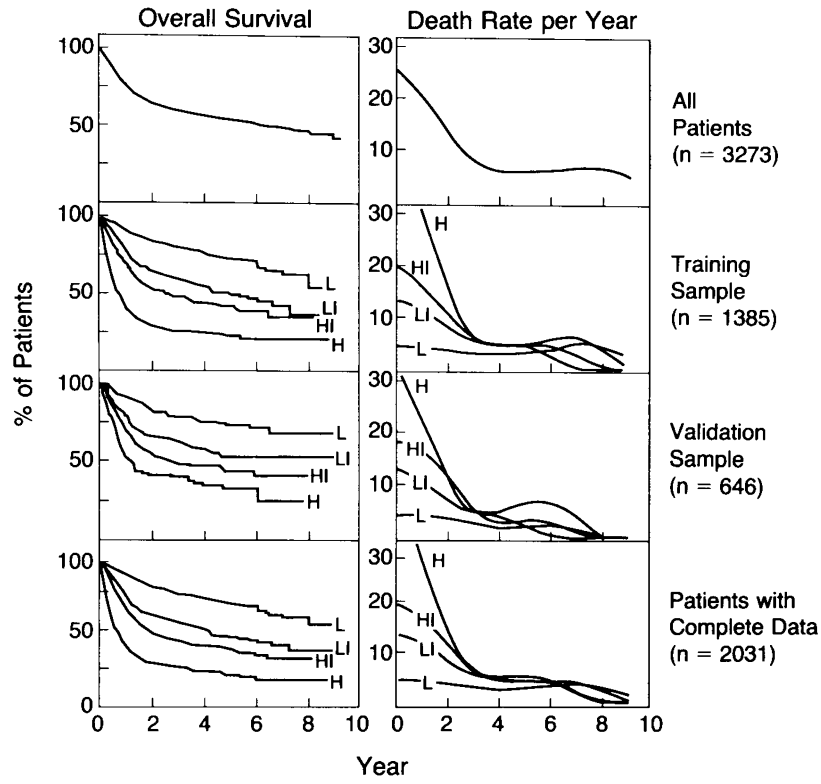

Figure 1. Survival According to Risk Group Defined by the International Index.
The left panels show Kaplan–Meier survival curves for the four risk groups (L denotes low risk, LI low intermediate risk, HI high intermediate risk, and H high risk). The right panels show death rates during the study period. Only 2031 of the 3273 patients had enough relevant information for classification according to the international index.

analyses (Table 3), a model incorporating these additional features would by definition be more predictive than the Ann Arbor classification system. This is illustrated in Figure 3, which shows the survival of patients in Ann Arbor stage II, III, and IV according to their risk group as defined by the international index.

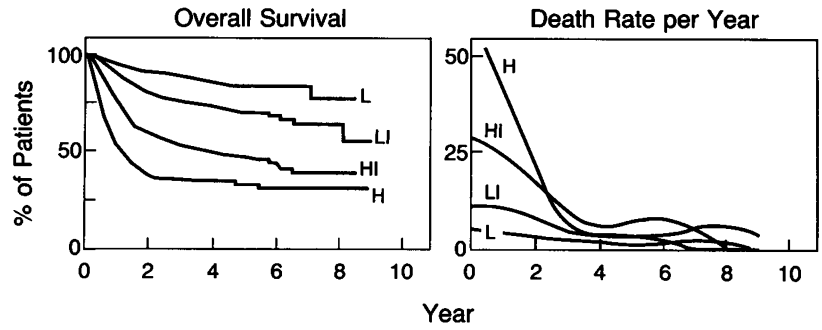

Figure 2. Survival among the 1274 Younger Patients (≤60 Years) According to Risk Group Defined by the Age-Adjusted International Index.

The left panel shows the Kaplan–Meier curves for this age group, and the right panel the death rates during the study period. L denotes low risk, LI low intermediate risk, HI high intermediate risk, and H high risk. Only 1274 of the 1931 patients 60 or younger had enough relevant information for classification according to the international index.

Discussion

The goal of our project was to develop a system for classifying patients with aggressive non-Hodgkin's lymphoma according to universally recognized clinical features. The resulting model applicable to all these patients (the international index) incorporates clinical features that reflect the growth and invasive potential of the tumor (tumor stage, serum LDH level, and number of extranodal disease sites), the patient's response to the tumor (performance status), and the patient's ability to tolerate intensive therapy (age and performance status). The simplified model for younger patients (the age-adjusted international index) uses a subgroup of these clinical features (tumor stage, LDH level, and performance status). Both models identified four risk groups of patients based on both the rate of complete response and the rate of relapse from complete response. The size of the study population and the diversity of the referring institutions and study centers helped ensure that the international index and the age-adjusted index were derived from a broadly representative group. Since recent studies indicate that unselected patients with aggressive non-Hodgkin's lymphoma who are treated with first-, second-, and third-generation chemotherapy regimens have comparable outcomes,[25-27] the variety of regimens that contained doxorubicin in our study is unlikely to have influenced the analysis. Furthermore, the international index was equally predictive of survival in two recent series of over 2000 patients treated with intensive third-generation regimens[25] (and unpublished data).

We developed separate models — one for all patients (the international index) and one for younger patients (the age-adjusted international index) — because the two models may be applicable in different settings. In trials that include patients of all ages, the international index, not restricted according to age, would be more useful. However, in trials of more intensive experimental approaches that are targeted to younger patients, the age-adjusted index could be used.

We retained four risk groups of patients defined by the international index and the age-adjusted index because physicians and investigators may collapse these risk groups differently, depending on their objectives. For example, if the goal is to compare the types of patients who are being treated in specific trials, the relative numbers of patients in all four risk groups should be

Table 4. Outcome According to Risk Group Defined by the International Index and the Age-Adjusted International Index.

Risk Group	No. of Risk Factors	Distribution of Patients (%)	Complete Response			Survival	
			Rate (%)	Relapse-Free Survival		2-yr Rate (%)	5-yr Rate (%)
				2-yr rate (%)	5-yr rate (%)		
International index, all patients (n = 2031)*							
Low	0 or 1	35	87	79	70	84	73
Low intermediate	2	27	67	66	50	66	51
High intermediate	3	22	55	59	49	54	43
High	4 or 5	16	44	58	40	34	26
Age-adjusted index, patients ≤60 (n = 1274)†							
Low	0	22	92	88	86	90	83
Low intermediate	1	32	78	74	66	79	69
High intermediate	2	32	57	62	53	59	46
High	3	14	46	61	58	37	32
Age-adjusted index, patients >60 (n = 761)†							
Low	0	18	91	75	46	80	56
Low intermediate	1	31	71	64	45	68	44
High intermediate	2	35	56	60	41	48	37
High	3	16	36	47	37	31	21

*The total of patients includes the 1385 in the training sample and the 646 in the validation sample.

†The total of the patients in the two analyses with the age-adjusted index includes four more patients than the total in the analysis with the international index because all the data necessary for these four patients to be included in the age-adjusted analyses (which evaluated fewer variables) were available.

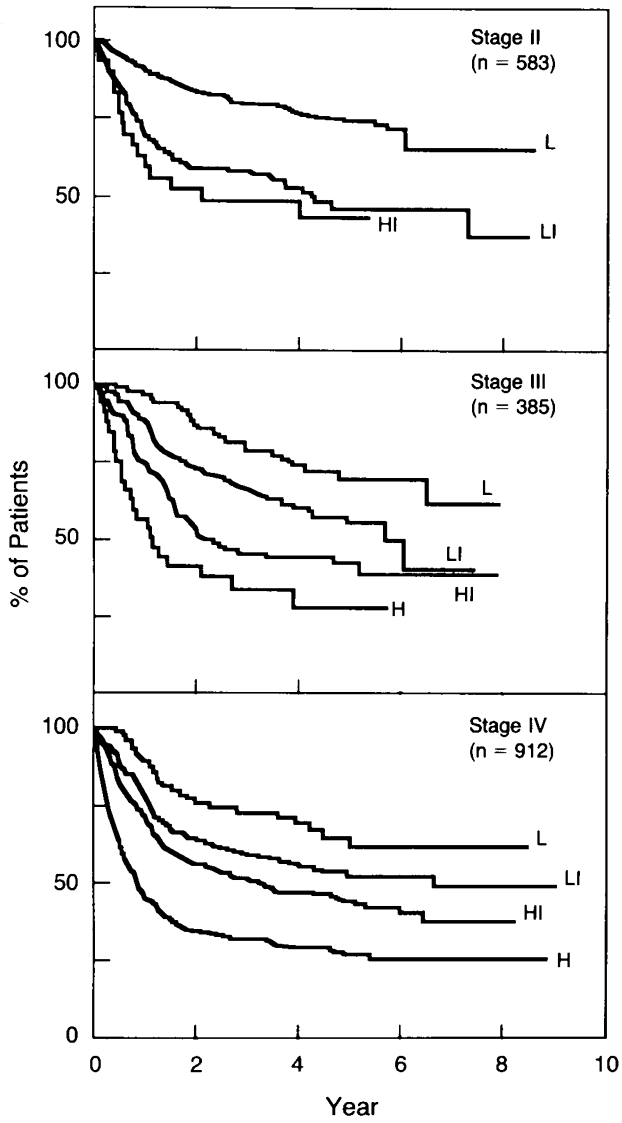

Figure 3. Survival among 1880 Patients in Ann Arbor Stages II, III, and IV, According to Risk Group Defined by the International Index.

L denotes low risk, LI low intermediate risk, HI high intermediate risk, and H high risk. Values for these groups are also shown in Table 4. Only one patient in stage II was at high risk, and that group is therefore not shown. Only 1881 of 3025 patients in stages II through IV had enough relevant information for classification according to the international index.

noted. If the objective is to identify candidates for experimental therapy — patients whose predicted five-year survival is less than 50 percent with standard regimens — it would be reasonable to include patients identified as being at high, high intermediate, and perhaps even low intermediate risk according to the international index (Table 4). However, when experimental approaches are specifically designed for younger patients (≤60 years), the target population might be patients at high intermediate and high risk as defined by the age-adjusted international index (Table 4). It is important that the target population of a "high-risk" protocol be accurately defined because the results of an experimental approach may be as dependent on the definition of high risk as on the regimen itself.[28,29] Furthermore, therapeutic approaches should be compared in appropriate, age-matched populations because younger patients generally have more favorable outcomes.

If it were possible to identify patients who enter complete remission but are at increased risk of subsequent relapse, such patients might be candidates for intensive experimental consolidation therapy. However, in our study, the clinical features that correlated with an increased risk of relapse were also associated with a decreased likelihood of obtaining an initial complete remission. Therefore, therapeutic approaches to patients at high risk must be directed toward increasing the low rates of initial complete responses as well as toward improving the durability of those responses.

Although the international index was specifically developed to predict outcome in patients with aggressive non-Hodgkin's lymphoma, it may also have prognostic value in patients with lymphoma that is histologically more indolent. Recent studies indicate that a prognostic-factor model developed for patients with aggressive non-Hodgkin's lymphoma also predicted survival in a small series of patients with follicular lymphoma,[30] and that the international index predicted survival in a larger series of similar patients (unpublished data).

Finally, it is important to recognize that the clinical prognostic features incorporated in the international index are, in part, surrogate variables that reflect the biologic heterogeneity of aggressive non-Hodgkin's lymphoma. As additional features such as serologic variables (β_2-microglobulin level[31]), indexes of tumor-cell proliferation (expression of Ki-67 antigen[32] and incorporation of tritiated thymidine[33]), karyotypic abnormalities,[34-37] and aberrant adhesion-molecule[38,39] and oncogene[40] expression are evaluated in larger numbers of patients, the biologic heterogeneity of this disease may be better understood. In the meantime, clinical prognostic-factor models such as the international index and the age-adjusted index can be used to identify specific risk groups and to compare different therapeutic approaches.

We are indebted to Jane Russell, Mia Donnelly, Stacey Cohen, and Sharon Cassidy of the Dana–Farber Cancer Institute Quality Control Center for assistance with data management; to Jane Weeks, M.D., for thoughtful review of the manuscript; and to Donna Favreau and Judy Nahas for assistance in the preparation of the manuscript.

References

1. Carbone PP, Kaplan HS, Musshoff K, Smithers DW, Tubiana M. Report of the Committee on Hodgkin's Disease Staging Classification. Cancer Res 1971;31:1860-1.
2. Rosenberg SA. Validity of the Ann Arbor staging classification for the non-Hodgkin's lymphomas. Cancer Treat Rep 1977;61:1023-7.

3. Fisher RI, Hubbard SM, DeVita VT, et al. Factors predicting long-term survival in diffuse mixed, histiocytic, or undifferentiated lymphoma. Blood 1981;58:45-51.
4. Shipp MA, Harrington DP, Klatt MM, et al. Identification of major prognostic subgroups of patients with large-cell lymphoma treated with m-BACOD or M-BACOD. Ann Intern Med 1986;104:757-65.
5. Jagannath S, Velasquez WS, Tucker SL, et al. Tumor burden assessment and its implication for a prognostic model in advanced diffuse large-cell lymphoma. J Clin Oncol 1986;4:859-65.
6. Danieu L, Wong G, Koziner B, Clarkson B. Predictive model for prognosis in advanced diffuse histiocytic lymphoma. Cancer Res 1986;46:5372-9.
7. Cowan RA, Jones M, Harris M, et al. Prognostic factors in high and intermediate grade non-Hodgkin's lymphoma. Br J Cancer 1989;59:276-82.
8. Coiffier B, Lepage E. Prognosis of aggressive lymphomas: a study of five prognostic models with patients included in the LNH-84 regimen. Blood 1989;74:558-64.
9. Coiffier B, Gisselbrecht C, Vose JM, et al. Prognostic factors in aggressive malignant lymphomas: description and validation of a prognostic index that could identify patients requiring a more intensive therapy. J Clin Oncol 1991;9:211-9.
10. Velasquez WS, Jagannath S, Tucker SL, et al. Risk classification as the basis for clinical staging of diffuse large-cell lymphoma derived from 10-year survival data. Blood 1989;74:551-7.
11. Hoskins PJ, Ng V, Spinelli JJ, Klimo P, Connors JM. Prognostic variables in patients with diffuse large-cell lymphoma treated with MACOP-B. J Clin Oncol 1991;9:220-6.
12. Dixon DO, Neilan B, Jones SE, et al. Effect of age on therapeutic outcome in advanced diffuse histiocytic lymphoma: the Southwest Oncology Group experience. J Clin Oncol 1986;4:295-305.
13. Shipp MA, Yeap BY, Harrington DP, et al. The m-BACOD combination chemotherapy regimen in large-cell lymphoma: analysis of the completed trial and comparison with the M-BACOD regimen. J Clin Oncol 1990;8:84-93.
14. Non-Hodgkin's Lymphoma Pathologic Classification Project. National Cancer Institute sponsored study of classifications of non-Hodgkin's lymphomas: summary and description of a working formulation for clinical usage. Cancer 1982;49:2112-35.
15. Lennert K, Mohri N, Stein H, Kaiserling E. The histopathology of malignant lymphoma. Br J Haematol 1975;31:Suppl:193-203.
16. Rappaport H. Tumors of the hematopoietic system. Sec. 3, Fasc. 8 of Atlas of tumor pathology. Washington, D.C.: Armed Forces Institute of Pathology, 1966.
17. Cox DR, Snell EJ. Analysis of binary data. 2nd ed. London: Chapman & Hall, 1989.
18. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc 1958;53:457-81.
19. Mantel N. Evaluation of survival data and two new rank order statistics arising in its consideration. Cancer Chemother Rep 1966;50:163-70.
20. Cox DR. Regression models and life-tables. J R Stat Soc [B] 1972;34:187-220.
21. Gray RJ. Some diagnostic methods for Cox regression models through hazard smoothing. Biometrics 1990;46:93-102.
22. Breiman L, Friedman JH, Olshen RA, Stone CJ. Classification and regression trees. Belmont, Calif.: Wadsworth, 1984.
23. Therneau TM, Grambsch PM, Fleming TR. Martingale-based residuals for survival models. Biometrika 1990;77:147-60.
24. U.S. decennial life tables for 1979-81. Vol. 1. No. 1. United States life tables. Hyattsville, Md.: Department of Health and Human Services, 1985. (DHHS publication no. (PHS) 85-1150-1.)
25. Gordon LI, Harrington D, Andersen J, et al. Comparison of a second-generation combination chemotherapeutic regimen (m-BACOD) with a standard regimen (CHOP) for advanced diffuse non-Hodgkin's lymphoma. N Engl J Med 1992;327:1342-9.
26. Fisher RI, Gaynor ER, Dahlberg S, et al. Comparison of a standard regimen (CHOP) with three intensive chemotherapy regimens for advanced non-Hodgkin's lymphoma. N Engl J Med 1993;328:1002-6.
27. Montserrat E, Garcia-Conde J, Viñolas N, et al. ProMACE-CytaBOM vs. CHOP in the treatment of unfavorable lymphomas: a randomized trial. Blood 1991;78:Suppl 1:127a. abstract.
28. Gulati SC, Shank B, Black P, et al. Autologous bone marrow transplantation for patients with poor-prognosis lymphoma. J Clin Oncol 1988;6:1303-13.
29. McMaster ML, Greer JP, Wolff SN, et al. Results of treatment with high intensity, brief duration chemotherapy in poor prognosis non-Hodgkin's lymphoma. Cancer 1991;68:233-41.
30. Bastion Y, Berger F, Bryon PA, Felman P, Ffrench M, Coiffier B. Follicular lymphomas: assessment of prognostic factors in 127 patients followed for 10 years. Ann Oncol 1991;2:Suppl 2:123-9.
31. Swan F Jr, Velasquez WS, Tucker S, et al. A new serologic staging system for large-cell lymphomas based on initial β2-microglobulin and lactate dehydrogenase levels. J Clin Oncol 1989;7:1518-27.
32. Grogan TM, Lippman SM, Spier CM, et al. Independent prognostic significance of a nuclear proliferation antigen in diffuse large cell lymphomas as determined by the monoclonal antibody Ki-67. Blood 1988;71:1157-60.
33. Cowan RA, Harris M, Jones M, Crowther D. DNA content in high and intermediate grade non-Hodgkin's lymphoma — prognostic significance and clinicopathological correlations. Br J Cancer 1989;60:904-10.
34. Cabanillas F, Pathak S, Grant G, et al. Refractoriness to chemotherapy and poor survival related to abnormalities of chromosomes 17 and 7 in lymphoma. Am J Med 1989;87:167-72.
35. Levine EG, Arthur DC, Frizzera G, Peterson BA, Hurd DD, Bloomfield CD. Cytogenetic abnormalities predict clinical outcome in non-Hodgkin lymphoma. Ann Intern Med 1988;108:14-20.
36. Schouten HC, Sanger WG, Weisenburger DD, Anderson J, Armitage JO. Chromosomal abnormalities in untreated patients with non-Hodgkin's lymphoma: associations with histology, clinical characteristics, and treatment outcome. Blood 1990;75:1841-7.
37. Yunis JJ, Mayer MG, Arnesen MA, Aeppli DP, Oken MM, Frizzera G. *bcl*-2 and other genomic alterations in the prognosis of large-cell lymphoma. N Engl J Med 1989;320:1047-54.
38. Jalkanen S, Joensuu H, Klemi P. Prognostic value of lymphocyte homing receptor and S phase fraction in non-Hodgkin's lymphoma. Blood 1990;75:1549-56.
39. Sy MS, Guo Y-J, Stamenkovic I. Distinct effects of two CD44 isoforms on tumor growth in vivo. J Exp Med 1991;174:859-66.
40. Zutter M, Hockenbery D, Silverman GA, Korsmeyer SJ. Immunolocalization of the Bcl-2 protein within hematopoietic neoplasms. Blood 1991;78:1062-8.