Comparison of Prognostic Models in Patients with Advanced Hodgkin Disease

Promising Results from Integration of the Best Three Systems

Paolo G. Gobbi, M.D.[1]
Pier Luigi Zinzani, M.D.[2]
Chiara Broglia, M.D.[1]
Mario Comelli, M.D.[3]
Massimo Magagnoli, M.D.[2]
Massimo Federico, M.D.[4]
Francesco Merli, M.D.[5]
Emilio Iannitto, M.D.[6]
Sante Tura, M.D.[2]
Edoardo Ascari, M.D.[1]

[1] Clinica Medica e Oncologia Medica, Università di Pavia, I.R.C.C.S. Policlinico S. Matteo, Pavia, Italy.

[2] Istituto di Ematologia e Oncologia Medica "L.e A. Seragnoli," Università di Bologna, Policlinico S. Orsola Malpighi, Bologna, Italy.

[3] Cattedra di Statistica, Istituto di Scienze Sanitarie e Statistiche, Università di Pavia, Pavia, Italy.

[4] Cattedra e Divisione di Oncologia Medica, Università di Modena, Modena, Italy.

[5] Servizio di Ematologia, Arcispedale "S. Maria Nuova," Reggio Emilia, Italy.

[6] Cattedra e Divisione di Ematologia, Università di Palermo, Palermo, Italy.

Supported in part by a grant from the Ministero per la Ricerca Scientifica e Tecnologica, Rome, and by the Fondazione "Adolfo Ferrata e Edoardo Storti," Pavia.

Address for reprints: Paolo G. Gobbi, M.D., Medicina Interna e Oncologia Medica, Università di Pavia, I.R.C.C.S. Policlinico S. Matteo, P. le Golgi no. 2, 27100 Pavia, Italy; Fax: 0039-0382-526223; E-mail: gobbipg@smatteo.pv.it.

Received June 8, 2000; revision received December 7, 2000; accepted January 2, 2001.

BACKGROUND. Several prognostic systems have been elaborated for patients with Hodgkin disease (HD) over the last 12 years, but early identification of a reasonably large group of both low and high risk, advanced stage patients remains unsatisfactory.

METHODS. Seven well known models were applied to 516 patients with advanced HD, with 315 patients used for the study sample and 201 patients used for the test sample. Individual performances as well as joint performances were analyzed univariately and multivariately in relation to overall survival, recurrence free survival, and time to treatment failure by means of a proportional hazards model.

RESULTS. None of the models identified a group containing $>$ 10% of patients from the total population who had a failure risk of either $\leq$ 10% or $\geq$ 50%. The systems of the International Database on Hodgkin Disease, the Memorial Sloan-Kettering Cancer Center, and the International Prognostic Factor Project showed the best prognostic power; only these three, when analyzed together, predicted clinical outcome with a statistically significant fit to the clinical data. Integration of the three systems in a linear model dramatically improved their individual discriminatory capacity by identifying patients with 10% and 50% failure risks, respectively, in 23% and 24% of the study patient population and in 19% and 25% of the test population, respectively.

CONCLUSIONS. As powerful and simple new prognostic factors are awaited that may improve our predictive ability, this integrated index is probably the best way to exploit the significance of those presently available. The program required for the calculations can be downloaded from the Internet at the web site http://www.unimo.it/gisl/default.htm. ***Cancer* 2001;91:1467–78.**

KEYWORDS: Hodgkin disease, prognosis, modeling, survival, time to treatment failure.

Defining different, well documented risk groups to differentiate the intensity of work-up and therapy is becoming increasingly important in patients with Hodgkin disease (HD). A large number of therapeutic strategies can be differentiated properly by taking into account the variable aggressiveness of initial clinical presentation, different staging accuracies, the probable effectiveness of salvage therapy in patients with disease recurrence, and the expected incidence of treatment-related late toxicity.[1,2] Studies of patient characteristics that demonstrate independent prognostic value try to define the tools for selecting patients in whom a durable complete remission is achievable with a reduction of therapy and related late toxicity or, more importantly, for selecting those in whom a high risk of unre-

sponsiveness or failure requires early intensification of therapy and acceptance of the correspondingly higher risk of adverse early and late consequences. Patients with advanced disease, as expected, represent one of the most important fields of clinical investigation in HD, because they have the greatest need to improve their cure rate, considering that about 33–50% of them ultimately succumb to the disease,[1,2] as well as the highest risk of late toxicity related to more aggressive treatment.

Several combinations of various prognostic factors have been proposed in the last decade,[3–9] and some of them have been challenged and compared[8,10,11] in series of patients with advanced HD, mainly for the purpose of identifying a sufficiently high risk group of patients for whom early intensive chemotherapy plus autologous stem cell transplantation (ASCT) or other investigational therapies could be recommended. At present, a less than 50% chance of being progression free at 5 years seems to be sufficient and appropriate to justify an early intensified therapeutic program;[12,13] however, the major drawback presented by each prognostic system is the small proportion of patients allocated to such a risk class—much lower than the 20–50% of patients with advanced disease who are destined to fail.[1,2,14] Conversely, the possible extension of early intensified therapy plus ASCT for patients with a less unfavorable prognosis would mean exposing a portion of patients who also may be cured with conventional therapies to the risks linked to such a treatment, not to mention an unjustified waste of resources and an increase of costs.

Moreover, not all of the available prognostic systems have been evaluated comparatively in previous investigations,[8,10,11] nor has clear and unequivocal information been given about which systems are relatively better or about how to improve their predictive power if possible. The current study tries to answer these questions, actually identifies three prognostic models that statistically fit the clinical data best, and suggests a technique for their combined use to improve accuracy in the selection of patients with substantially different failure risks.

MATERIALS AND METHODS

The analysis was performed on a population of 315 patients with advanced HD who were collected by combining 145 patients who were treated at the Italian Lymphoma Study Group (GISL) and 170 patients who were studied at the Institute of Hematology of Bologna (IHB). These patients, whose characteristics are detailed in Table 1, shared nearly identical overall survival (OS), recurrence free survival (RFS), and time to treatment failure (TTF) rates and could be pooled into

TABLE 1
Clinical Characteristics of the Two Patient Series (Italian Lymphoma Study Group and Institute of Hematology of Bologna) that Entered the Study Sample and of the Patient Population Taken as Test Sample (Italian Leukemia Association)

	GISL		IHB		AIL	
Characteristic	**No.**	**%**	**No.**	**%**	**No.**	**%**
Patients	145	100	170	100	201	100
Gender						
Male	85	59	88	52	118	59
Female	60	41	82	48	83	41
Age (yrs)						
≤ 14	0	0	4	2	0	0
15–19	17	12	31	18	17	9
20–29	43	29	56	33	70	35
30–39	27	19	34	20	48	24
40–49	27	19	23	13	24	12
50–59	10	7	16	9	23	11
60–69	17	11	6	4	16	7
70–75	4	3	0	0	3	1
Histology						
Lymphocyte predominance	2	1	5	3	9	4
Nodular sclerosis	88	60	140	82	118	59
Mixed cellularity	42	30	20	12	62	31
Lymphocyte depletion	13	9	5	3	12	6
Stage IA	—	—	5	3	—	—
IB	—	—	8	5	—	—
IIA	—	—	25	15	—	—
IIB	45	31	18	11	48	24
IIIA	22	15	30	18	38	19
IIIB	39	27	55	32	66	33
IVA	11	8	14	8	13	06
IVB	28	19	15	9	36	18
Mediastinal bulk	47	32	68	40	51	25
Bone marrow involvement	22	15	8	5	18	9
Hemoglobin < 10 g/dL	16	11	12	7	35	17
ESR > 40 mm/1st hr	95	66	66	39	127	63
Serum LDH > 450 mU/mL	33	23	2	1	36	18
Serum albumin < 3 g/dL	26	18	3	2	34	17
Response						
Complete	137	94	152	89	167	84
Partial	4	3	1	< 1	14	7
Null	4	3	9	5	11	5
Progression	0	0	5	3	6	3
Recurrences after CR	24	17	26	15	47	23
Death due to HD	21	15	36	21	53	25
Chemotherapy (no. of cycles)						
MOPP/EBV/CAD	6	—	—	—	—	—
MOPP/ABVD	—	—	8	—	—	—
CVPP (or MOPP)/AVBD	—	—	—	—	12	—
Radiotherapy (optional)[a]	48	33	101	59	134	67
Median follow-up (months)	81	—	109	—	139	—
P value						
OS at 5 yrs and 10 yrs	0.87/0.82		0.86/0.77		0.80/0.68	
RFS at 5 yrs and 10 yrs	0.83/0.79		0.89/0.76		0.81/0.70	
TTF at 5 yrs and 10 yrs	0.79/0.78		0.77/0.70		0.69/0.61	

GISL: Italian Lymphoma Study Group; IHB: Institute of Hematology of Bologna; AIL: Italian Leukemia Association; ESR: erythrocyte sedimentation rate; LDH: serum lactate dehydrogenase; CR: complete response; HD: Hodgkin disease; MOPP: mechlorethamine, vincristine, procarbazine, and prednisone; OS: overall survival; RFS: recurrence free survival; TTF: time to treatment failure; EBV: epidoxorubicin, bleomycin, vinblastine; CAD: CCNU, melphalan, vindesine; CVPP: CCNU, vinblastine, procarbazine, prednisone.

[a] Radiotherapy was delivered only to a few sites with tumor masses that had either been particularly large at presentation or reduced slowly or partially during chemotherapy.

a single study sample population. Table 1 also illustrates the clinical features of the 201 patients with advanced HD from the Italian Leukemia Association (AIL) who showed similar, albeit slightly lower and not significantly different, results and who were used as an external test sample.

Enrollment criteria, clinical presentation features, staging procedures, and treatment for the whole GISL population[15,16] and a large majority of the IHB[17] and AIL[18] series were described elsewhere. The GISL trial accrued patients from January 1, 1988 to September 30, 1993; the IHB series accrued patients between January 15, 1979 and December 15, 1994; and the AIL series accrued patients between January 1, 1980 and December 31, 1985. Pathology was reviewed within each group. Staging criteria routinely included computed tomography scans of the thorax and abdomen and unilateral bone marrow biopsy. Furthermore, most patients underwent abdominal ultrasonography. None of the GISL patients, 12 of the IHB patients, and 7 of the AIL patients were staged through exploratory laparotomy with splenectomy. All three series collected all patients with Stage III or IV disease. In addition, the GISL and AIL series included Stage II patients with B symptoms according to their unfavorable prognostic estimate,[20] and the IHB enrolled patients with early stage disease (I and II) who presented with large mediastinal masses.[20,21] The management of patients with early stage disease with one or more of these unfavorable prognostic factors is similar to that of patients with advanced stage disease, in agreement with current statements and practice.[20,23]

The GISL patients were treated with 6 cycles of mechlorethamine, vincristine, and procarbazine (MOPP)/EBV/CAD chemotherapy, the IHB patients received an alternating MOPP/ABVD regimen for 8 cycles, and the AIL patients were randomized to receive 12 cycles of either CVPP/ABVD or MOPP/ABVD alternating multiple drug therapy.[18] In the GISL and IHB series, optional radiotherapy (RT) was delivered after chemotherapy only to some of the lymph node areas with previous major involvement or to those areas that were reduced partially or more slowly during chemotherapy. Thirty-three percent of GISL patients had combined RT, the recommended total dose of which did not exceeded 35 grays (Gy). Most of the IHB patients (59%) received RT limited to the involved lymph node areas after the end of chemotherapy, with a total dose that generally did not exceed 36 Gy, whereas 62% of the AIL patients underwent RT, which was administered to extended fields in patients with Stage IIB and III disease and to selected involved fields in patients with Stage IV disease.

Complete remission (CR) was defined as the complete regression of measured lesions and the disappearance of all other objective evidence of disease for at least 3 months. Partial remission (PR) consisted of a decrease $>$ 50% in the sum of the products of the greatest dimensions of measurable lesions with resolution of symptoms, if present at onset. No response (NR) was defined as a variation in the sum of the products of measurable lesions ranging from a $<$ 50% decrease up to a 25% increase. Disease was considered as progressive (PD) when there was a $>$ 25% increase in the size of at least one measurable lesion.[24]

OS was calculated from the date of diagnosis to the date of the last observation or death. RFS for complete responders was measured from the date of therapy completion to the date of the last observation or disease recurrence. TTF was computed from the start of treatment to one of the following events: disease progression during treatment, no CR at the end of treatment, disease recurrence, or death from disease. PR and NR were considered as events, because they often hide a resistant or refractory tumor component,[25] at least as early recurrences ($<$ 12 months). These generally are considered unquestionable events, and patients with incomplete remissions share most of their therapeutic requirements.[26] The advantage of using TTF data for this study is that they pool all types of failures that can be related to unsuccessful therapy and for which alternative or early intensified treatments[22,23] may be justified: incomplete response (PR and NR), disease progression, disease recurrence, death from the disease. Deaths due to causes other than HD were censored for RFS and TTF calculations.[27] Curves were calculated using the method of Kaplan and Meier.[28]

Seven prognostic models formulated for clinical purposes in HD over the last 10 years were taken into account and tested for their predictive power on the entire population of 315 patients. The prognostic models analyzed were the following: the equation described at the Universities of Pavia and Modena (PV-MO);[3] the index derived from the series of St. Bartholomew's Hospital and Christie Hospital (SB-C);[4] the list of factors identified at the Memorial Sloan-Kettering Cancer Center (MSK);[5] the index proposed by the Scottish and Newcastle Lymphoma Group (SNLG);[6] the model drawn from the International Database on Hodgkin's Disease (IDHD),[7] which was used here through the standardized deviate of the probability of surviving for 10 years; the Manchester Lymphoma Group (MLG)[8] index; and, finally, the index elaborated from the International Prognostic Factors Project (IPFP).[9] Table 2 illustrates the main characteristics of the prognostic indices analyzed. It is possible that patients who were studied previously by the PV-

TABLE 2
Main Characteristics of the Seven Prognostic Models Studied

Index	Clinical and laboratory parameters utilized
PV-MO (Gobbi et al. 1988[3]): OS parametric model (exponential); study, 586 patients; control, 179 patients	(329.5 − 64.6 × ESR − 70.6 × St − 60.2 × Hist − 40.4 × age − 29.9 × Alb − 24.3 × Gender) × 0.693 = expected median survival time in months
SB-C (Wagstaff et al. 1988[4]): OS nonparametric model (proportional hazard); study, 301 patients (St IIIB–IV only)	Low risk: (age < 45 yrs + Ly > 0.75 × 10^9/L) or (female + St IIIB); intermediate risk: rest; high risk: male + St IV + (age > 45 yrs or Ly < 0.75 × 10^9/L)
MSK (Straus et al. 1990[5]); OS and FFP nonparametric models (proportional hazard); study, 185 patients	Groups of increasingly severe prognosis according to the number of the following unfavorable factors: age > 45 yrs, LDH > 400 U/L, Ht < 38% (34% for females), inguinal involvement, mediastinal bulk > 0.45 for OS; the same five factors plus bone marrow involvement for FFP
SNLG (Proctor et al. 1991[6]): OS nonparametric model (proportional hazard); study, 92 patients; control, 455 patients	1.5858 − 0.0363 × age + 0.0005 × age^2 + 0.0683 × St − 0.086 × Ly − 0.0587 × Hb + 0.3 × bulk = adimensional prognostic index (risk: < 0, low; 0–0.3, low intermediate; 0.3–0.5, high intermediate; > 0.5, = high)
IDHD (Gobbi et al. 1994[7]): OS parametric model (log-normal); study, 2542 patients; control, 2481 patients	Exp [3.75 + 1.25 × St. I + 0.77 × St. II + 0.46 × St. III − 0.00046 × age^2 + 0.85 × Hist + 0.42 × B symptoms + ln(Alb distribution percentile) + 0.25 × Gender + 0.25 × IAD] = expected median survival time in months, probability of surviving a given time
MLG (Lee et al. 1997[8]): FFP nonparametric model (multiple hazard); study, 453 patients (St II–IV)	Low risk: St III–IV, neither bulk nor B symptoms; low-intermediate risk: St II + (bulk or B symptoms); high-intermediate risk: St III–IV + (bulk ± B symptoms) + Ly > 0.6 × 10^9/L; high risk: St III–IV + (bulk ± B symptoms) + (Ly ≤ 0.6 × 10^9/L ± involved marrow)
IPFP (Hasenclever and Diehl, 1998[9]): FFP nonparametric model (proportional hazard); study, 5141 patients	Groups of increasingly severe prognosis according to the number of the following unfavorable factors: Alb < 4 g/dL, Hb < 10.5 g/dL, male gender, St IV, age > 45 yrs, WBC ≥ 15 × 10^9/L, Ly < 0.6 × 10^9/L (or < 8% of WBC)

PV-MO: Universities of Pavia and Modena; OS: overall survival; ESR: erythrocyte sedimentation rate (mm at 1st hour); St: stage; Hist: histology (four types); Alb: serum albumin concentration (g/dL); SB-C: St. Bartholomew's Hospital and Christie Hospital; Ly: lymphocyte count (peripheral blood); MSK: Memorial Sloan-Kettering Cancer Center; FFP: freedom from progression; LDH: serum lactate dehydrogenase concentration (mU/mL); Ht: hematocrit (%); SNLG: Scottish and Newcastle Lymphoma group; Hb: hemoglobin (g/dL); IDHD: International Database on Hodgkin's Disease; IAD: involved area distribution (≤ 3 above the diaphragm or otherwise); MLG: Manchester Lymphoma Group; IPFP: International Prognostic Factors Project; WBC: white blood cell count.

MO and SB-C groups had been included in the IDHD, but the number of shared patients cannot be greater than 10% of the 5023 patients from the IDHD. Four models (PV-MO, SB-C, SNLG, and IDHD) considered only OS as the time dependent variable, and two other models (MLG and IPFP) took into account only freedom from progression (FFP), whereas the MSK system used both OS and FFP. Two indices were drawn from parametric models after they were demonstrated to fit the original data well. The remaining five indices emerged from semiparametric models (proportional hazards models). The selection procedures from which the predictive factors of each index were identified were similar, because they consisted of multiple regression techniques in which the goodness-of-fit measure was the likelihood of the model. The parametric models allow direct estimates of survival in terms of time units or the probability of surviving at a given time, whereas the semiparametric models lead to categorizations of a proper number of distinct, contiguous ranges of hazards (or survival probabilities) to differentiate prognosis for clinical uses.

All prognostic models were computed successfully in both the study and test sample populations according to the respective authors' recommendations, and this was possible for all patients (there were no missing values among the required data), with the following two minor modifications: bulky mediastinal mass entering the MSK system had to be accepted as having a ratio of ≥ 0.33 instead of ≥ 0.45 between the width of the enlarged mediastinum and that of the chest, because only the former information had been recorded and computerized. Second, the few patients with early stage disease who had an unfavorable presentation in the current series, who cannot be allocated into the original defining characteristics of the SB-C system (which considers Stage III–IV patients only) or of the MLG index (which takes into account only Stage II patients with B symptoms or bulky mediastinum; see Table 2), were allocated to distinct new risk groups "0" to allow a complete comparison of all indices on the whole patient population.

The ability of each prognostic index or of the combination of the best of them to predict the failure times was investigated in the study sample through univariable and multivariable proportional hazards models.[29] Both backward and forward selection techniques were used to explore the interrelations existing among models and to identify those with more independent predictive power. Finally, the results ob-

TABLE 3
Goodness of Fit of the Seven Prognostic Indexes to the Distribution of Overall Survival, Recurrence Free Survival, and Time to Treatment Failure Data from 315 Patients with Advanced Hodgkin Disease

OS			RFS			TTF		
Index	Chi-square of the likelihood ratio	*P* value	Index	Chi-square of the likelihood ratio	*P* value	Index	Chi-square of the likelihood ratio	*P* value
IDHD	40.282	< 0.0001	IDHD	28.618	< 0.0001	IDHD	34.076	< 0.0001
MSK (OS)	31.561	< 0.0001	MSK (FFP)	22.314	< 0.0001	MSK (FFP)	33.556	< 0.0001
SB-C	16.427	< 0.0001	PV-MO	17.458	< 0.0001	PV-MO	18.897	< 0.0001
PV-MO	16.312	< 0.0001	MSK (OS)	17.423	< 0.0001	SB-C	14.741	0.0001
IPFP	15.581	< 0.0001	SB-C	16.868	< 0.0001	SNLG	9.890	0.0017
SNLG	11.913	0.0006	SNLG	12.065	0.0005	IPFP	8.526	0.0035
MLG	4.235	0.0269	MLG	12.062	0.0005	MLG	7.974	0.0047
—	—	—	IPFP	5.359	0.0206	—	—	—

OS: overall survival; RFS: recurrence free survival; TTF: time to treatment failure; IDHD: International Database on Hodgkin's Disease; MSK: Memorial Sloan-Kettering Cancer Center; FFP: freedom from (disease) progression; SB-C: St. Bartholomew's Hospital and Christie Hospital; PV-MO: Universities of Pavia and Modena; IPFP: International Prognostic Factors Project; SNLG: Scottish and Newcastle Lymphoma Group; MLG: Manchester Lymphoma Group.

tained in the study sample were checked and validated in the test sample population.

RESULTS

Table 3 shows how accurately each prognostic system actually fits the distribution of OS, RFS, and TTF data for the 315 patients from the study sample (GISL and IHB series). The IDHD and MSK systems showed relatively better correspondence to the clinical evolution of the patient population, irrespective of the time parameter considered and that for which each index had been specifically devised (only the MSK system had a distinct formulation for OS and FFP).

Considering TTF as the best time dependent parameter for evaluating the risk of not being cured by conventional therapy, Figure 1 illustrates the proportions of patients with advanced HD allocated to the groups codified by each index and the corresponding failure risks observed. For models characterized by continuous distribution of values (IDHD, SNLG, and PV-MO), the categorized ranges suggested by the respective authors were used. The first four systems (IDHD, MSK, IPFP, and PV-MO) in the figure seem to use their clinical parameters most adequately, and they present prognostic classes of expected increasing clinical severity that correspond to increasing actual risk of failure. Moreover, the first three models are the most well balanced with respect to patient allocation to their prognostic groups, even though the strong difference in patient frequency between intermediate risk groups and extreme risk groups remains a problem. For example, the MSK system actually does identify patients with a 100% risk of failure, but this selection is made in only 0.6% of the total population. Alternatively, it can assign a failure risk $\geq$ 55% to 9.2% of all patients (Groups 3 and 4). The IPFP index finds only 9.5% of patients with a score $\geq$ 4 instead of the expected 19% of the original study, whereas the corresponding risk is 43%: acceptably near the expected 47%. Conversely, the IDHD and PV-MO models seem to be able to select patients with a failure risk $\leq$ 10%, but they do so in 9.5% and 12.5% of the overall population, respectively. The group "0" of the SB-C and MLG prognostic systems, assigned here to collect the unfavorable early stage patients who were not included in their original criteria, proves to be a suitable choice for risk grading and patient selection, but it does not help in identifying a large enough group with a favorable prognosis. In summary, Figure 1 clearly demonstrates that each index is able to identify only small numbers of patients with either a very low risk or an undoubtedly high clinical risk.

To make a closer comparison of these prognostic systems and to explore their possible interrelations, a multivariable analysis was performed with a proportional hazards model using TTF as the time dependent variable. Table 4 demonstrates that three systems carry nearly the total amount of predictive information: the IDHD, MSK, and IPFP models, making the contributions of the remaining four systems quite negligible. Moreover, around the pivotal role of the IDHD model, a remarkable quantity of additional independent information is provided by the MSK index, whereas the IPFP model is able to add only a limited, although statistically significant, predictive contribution.

It appeared fully logical, in the absence of prognostically more powerful single factors or multiple

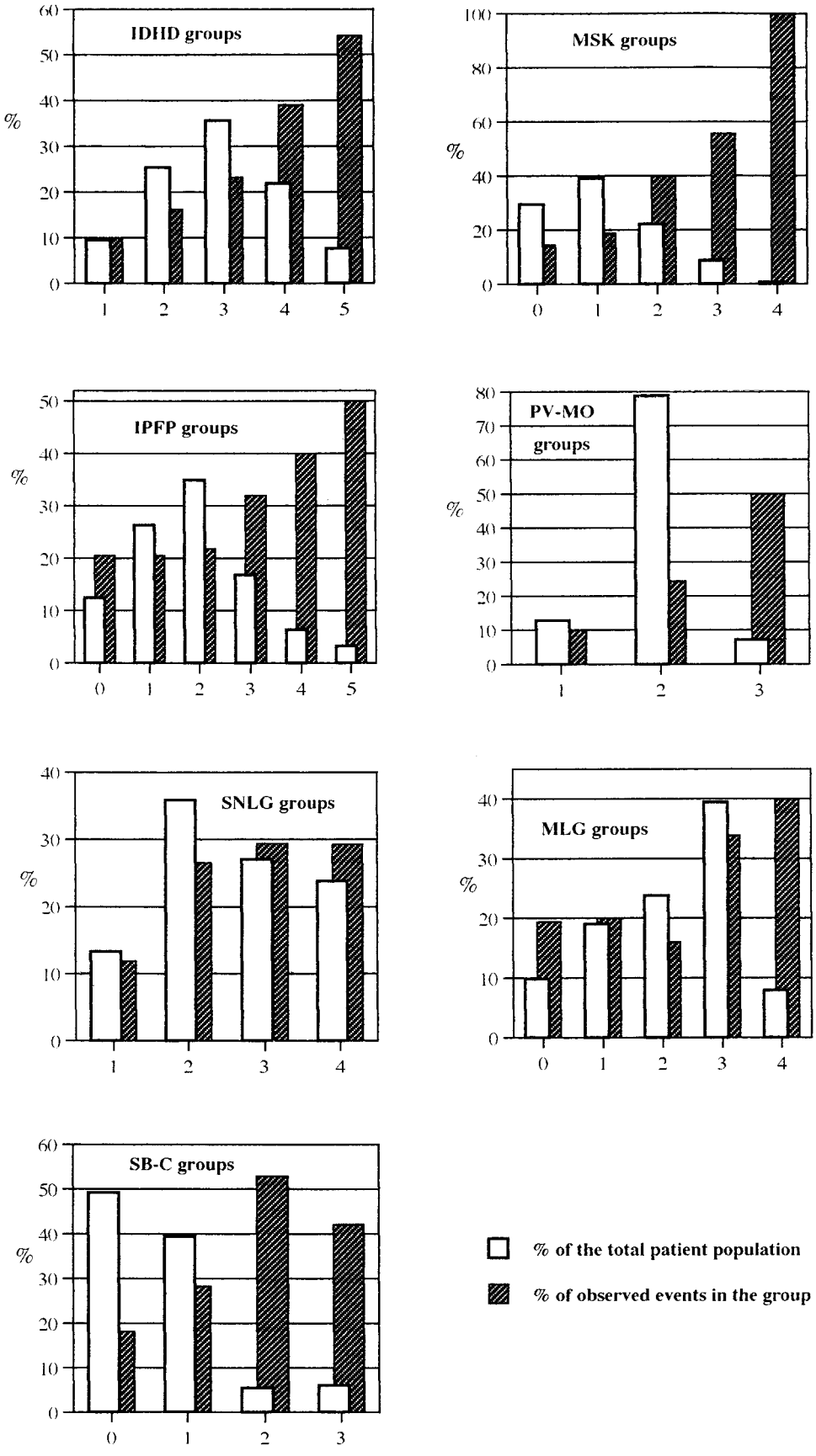

FIGURE 1. Comparison of the proportion of the patient population identified by the categories of each prognostic model together with an evaluation of the percentage risk of failure observed for each group identified. The groups in the International Database on Hodgkin's Disease (IDHD), Universities of Pavia and Modena (PV-MO), and Scottish and Newcastle Lymphoma Group (SNLG) models derive from categorization of their continuously distributed indexes according to the following ranges suggested in the original works: IDHD (probability of surviving at 120 months): 1, $>$ 0.850; 2, 0.850 $\div$ 0.750; 3, 0.749 $\div$ 0.590; 4, 0.589 $\div$ 0.360; 5, $<$ 0.360 (corresponding, respectively, to the given cut-off limits of expected survival at 120 months[25]: 1, $>$ 723; 2, 723 $\div$ 377; 3, 376.9 $\div$ 178; 4, 177.9 $\div$ 64; 5, $<$ 64 months). PV-MO cut-off values of expected survival at 120 months: 1, $>$ 200; 2, 199 $\div$ 73; 3, $<$ 72 months. SNLG cut-off values: 1, $<$ 0; 2, 0 $\div$ 0.299; 3, 0.3 $\div$ 0.5; 4, $>$ 0.5. MSK: Memorial Sloan-Kettering Cancer Center; IPFP: International Prognostic Factors Project; MLG: Manchester Lymphoma Group; SB-C: St. Bartholomew's Hospital and Christie Hospital.

TABLE 4
Multivariate Analysis of All Prognostic Models with Time to Treatment Failure as the Dependent Variable (Forward and Backward Selection)[a]

Index	Coefficient	Chi-square of the likelihood ratio	*P* value
IDHD	0.950	16.122	< 0.0001
MSK (FFP)	0.541	13.061	0.0003
IPFP	−0.302	5.487	0.0192
SNLG	—	2.369	0.1237
PV-MO	—	1.270	0.2598
MLG	—	0.003	0.9566
SB-C	—	0.066	0.7972

IDHD: International Database on Hodgkin's Disease; MSK (FFP): Memorial Sloan-Kettering Cancer Center (freedom from progression); IPFP: International Prognostic Factors Project; SNLG: Scottish and Newcastle Lymphoma Group; PV-MO: Universities of Pavia and Modena; MLG: Manchester Lymphoma Group; SB-C: St. Bartholomew's Hospital and Christie Hospital.

[a] Integrated index = 0.950 × IDHD (as probability of surviving at 120 months) + 0.541 × MSK groups (from 0 to 5) − 0.302 × IPFP groups (from 0 to 7).

systems, to try to integrate the independent information given by each of the three models in the hope of increasing their individual prognostic ability. All three models, with the regression coefficients derived from the preceding multivariable analysis, were entered into a linear expression (integrated index = 0.950 × IDHD [as the probability of surviving at 120 months] + 0.541 × MSK groups [0–5] − 0.302 × IPFP groups [0–7]) with values that were put in relation to the observed TTF. Figure 2 illustrates the correlations among the values of this integrated index, the corresponding cumulative failure risk, and the proportion of patients identified by index values. Two curves are obtained: the first correlates cumulative risk and the percent of patients presenting an integrated index under a given value, which better estimates the favorable good-risk tail of the population of patients with advanced HD, and the second correlates cumulative risk and the percent of patients showing an integrated index over a given value, which better studies the patients with an unfavorable prognosis. This figure can be used as a nomogram to select given index values with corresponding cumulative failure risk and expected proportion of involved patients (see Additional Information). Such an integrated index, despite its complexity, may represent the best refinement possible for evaluating the prognostic factors available at this time. In fact, Figure 2 shows that the index is able to identify a good fraction (23%) of the study sample patients who, by virtue of a value ≤ 0.90, show a low risk (≈ 10–15%). Conversely, a relatively large proportion of patients (24% of the total) with an index > 0.05 demonstrate a high risk of failure (≥ 50%). Each of these prognostically opposite groups, with clearly different therapeutic requirements, seems to be about twice as numerous as the groups identified by the best prognostic systems elaborated during the last 10 years, as shown in Figure 1.

Table 5 checks the predictive accuracy of the integrated index in the test sample of the 201 AIL patients and also allows a new comparison with the three individual systems from which it was derived. First, the integrated index proved to be capable of selecting from the whole set of patients, on the one hand, 24% with a low risk (≤ 14%) of unfavorable events and, on the other hand, 25% with a risk > 50%. None of the other prognostic systems identified similar proportions of patients with either a comparably favorable prognosis (IDHD: 30% with a risk ≤ 18%; MSK: 26% with a risk ≤ 19%; IPFP: 6% with a risk ≤ 15%) or a comparably unfavorable prognosis (IDHD: 11% with a risk ≥ 61%; MSK: 12% with a risk ≥ 57%; IPFP: 13% with a risk ≥ 57%). The table also shows the subsets of patients who were allocated incorrectly into the groups of low, intermediate, and high risk by each prognostic model compared with the allocation allowed by the integrated index with its better predictive accuracy.

This may mean that a small but interesting further step has been taken in the early identification of those patients who are known to respond either very well or very poorly to conventional therapy. Figure 3 is an example of the different index values that may be chosen to select a number of possible combinations of values with a given risk and the corresponding fraction of patients presenting that risk. From this point of view, this integrated index may help in selecting patients for investigational treatments and in making wiser use of technical and financial resources in the clinical management of patients.

DISCUSSION

HD has been an oncologic model from which several clinical achievements have been obtained. Whereas the evolution of RT equipment and radiation techniques allowed identification of the tumoricidal dose for HD, which cannot be increased further without unavoidable toxicity to normal tissues,[30] the well known responsiveness of the disease to chemotherapy and the correlation (first demonstrated in HD[31]) between drug dose and clinical response have encouraged the delivery of higher dosages, up to and exceeding the limit of myelotoxicity, thanks to the possible rescue offered by ASCT. This myeloablative intensification is generally applied after a first recurrence[23] or, with poorer results, after unsuccessful first-line chemotherapy; however, several investigators tend to propose similar therapeutic procedures up front in high

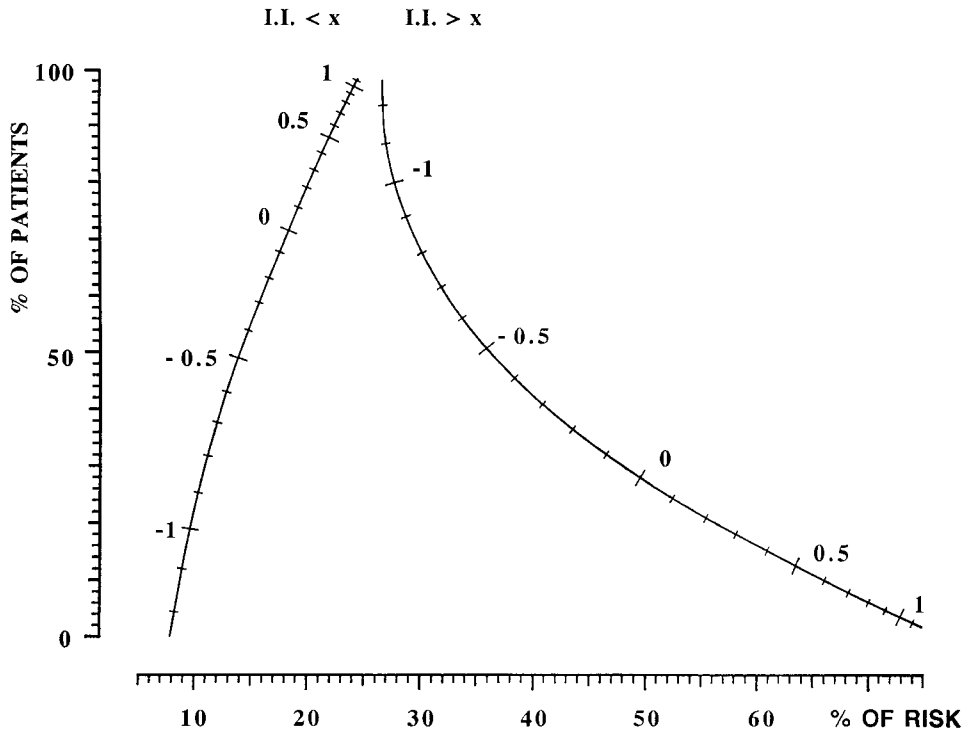

FIGURE 2. Correlations among values of the integrated index (I.I.), cumulative failure risk (observed over a median follow-up of 96 months), and proportion of patients with advanced Hodgkin disease who presented an I.I. either less than or greater than a given value. A number of combinations of possible risk values and corresponding percentages of the patient population can be sorted according to investigational purposes, therapeutic options, or financial resources.

risk, responding patients at the time of diagnosis:[32] In this condition, tolerance is expected to be better, and disease is anticipated to be less resistant. Nevertheless, to avoid exposing patients who are potentially curable with conventional therapy to transplantation-related risks, an effective prognostic tool is needed to single out early those patients in whom the risks of intensification are justified by intrinsic disease hazard. Similar reasoning holds true for investigational therapies other than those requiring ASCT that are proposed in the same way for high risk patients, such as adoptive cellular immunotherapy or radiolabeled or toxin-conjugated monoclonal antibodies. Moreover, high cost and complexity are additional, albeit secondary arguments shared by all nonconventional therapies in favor of an accurate prognostic selection of their potential candidates.

Thus, investigation of prognostic factors in patients with HD is still fully justified, and the large body of work conducted in recent years has led to the definition and validation of a number of sets of prognostic systems, often differing somewhat regarding the number and type of parameters involved, the weight given to each of them, the patient series from which they were extracted, the clinical purposes pursued, and the techniques of analysis adopted. Reproducibility of results among centers has been a constant problem in these investigations. In this regard, even ironic remarks about "magical formulas" have had to be tolerated from clinicians.[33]

However, some common features can be found in the information collected over the last 10 years. First, differences in the time dependent variable used (OS, RFS, FFP, FFS, etc.)—although they disrupt comparative evaluations—are not crucial and probably are not responsible for the selection of different prognostic variables in the models designed so far. Table 3 demonstrates that, no matter which survival parameter has been used in each prognostic index, the rank of the model's predictive capacity is roughly respected when considering OS, RFS, or TTF. This is likely because HD is the main common determinator of the events considered by each survival time parameter (progression, recurrence, and death). However, we agree with Carde's opinion[10] that, when evaluating prognosis with the specific aim of guiding treatment options, the most proper time dependent variable is the one that, like TTF in the current work, also considers partial and null remissions as events as well as those included in FFP, i.e., progression during therapy, recurrence at any time, and death from disease. Thus, in TTF, all possible events depending on unsuccessful therapy are included, both failure to achieve CR and failure to maintain it. Furthermore, follow-up length is at least as important as the choice of a particular time dependent criterion for evaluating results. A 3-year median follow-up can be considered rather short, as acknowledged by some authors themselves,[11] when comparing prognostic models. Eighteen months from the start of therapy may be too short a follow-up, espe-

TABLE 5
Proportion of the Test Patient Population Identified by Classes, Scores, or Ranges of the Original Models and by the Integrated Index and the Risk of Event Recorded in Each Patient Group Identified[a]

	Integrated index ranges			
No. of patients	< − 0.9 (49)	−0.9–0.05 (102)	> 0.05 (50)	Total (201)
Integrated index	0.24 (0.14)	0.51 (0.32)	0.25 (0.58)	1.0 (0.34)
IDHD groups				
I	0.01 (0.00)	—	—	0.01 (0.00)
II	0.15 (0.13)	0.13 (0.22)	—	0.29 (0.18)
III	0.08 (0.18)	0.20 (0.29)	0.05 (0.40)	0.34 (0.28)
IV	< 0.01 (0.00)	0.14 (0.38)	0.10 (0.67)	0.25 (0.49)
V	—	0.02 (0.75)	0.09 (0.58)	0.11 (0.61)
MSK score				
0	0.19 (0.13)	0.07 (0.13)	—	0.26 (0.19)
1	0.06 (0.17)	0.32 (0.34)	0.01 (0.50)	0.39 (0.32)
2	—	0.11 (0.22)	0.11 (0.57)	0.23 (0.39)
3	—	—	0.11 (0.57)	0.11 (0.57)
4	—	—	0.01 (1.0)	0.01 (1.0)
5	—	—	—	—
IPFP score				
0	0.01 (0.00)	0.05 (0.20)	—	0.06 (0.15)
1	0.10 (0.20)	0.11 (0.27)	0.01 (0.50)	0.23 (0.25)
2	0.09 (0.11)	0.16 (0.36)	0.07 (0.40)	0.33 (0.30)
3	0.03 (0.17)	0.12 (0.25)	0.06 (0.77)	0.21 (0.40)
4	< 0.01 (0.00)	0.04 (0.50)	0.08 (0.63)	0.12 (0.56)
5	< 0.01 (0.00)	0.02 (0.50)	0.01 (0.33)	0.04 (0.38)
6	—	—	< 0.01 (1.0)	< 0.01 (1.0)

IDHD: International Database on Hodgkin's Disease; MSK: Memorial Sloan-Kettering Cancer Center; IPFP: International Prognostic Factors Project.
[a] Numbers in parentheses indicate the risk of event.

cially if not only events beyond this limit but also those during therapy are not considered.[8] In this regard, it is well known that late recurrences can be cured with conventional chemotherapy, but it appears to be suitable not to exclude them from unfavorable events, because the problem is whether such recurrences can be predicted early, and, thus, first-line therapy can take them into account, and the need for and risk of a second treatment can be avoided.

Second, both the number and the heterogeneity of predictive variables identified by prognostic models have been considered as a possible confirmation of their low reliability,[11,33] although this may be only a minor problem. In fact, if we examine the list of all parameters included in the systems from the last 12 years (Table 2), then we can see that, from a total of 12,415 patients studied, the 7 models selected no more than 15 prognostic variables in all. The factors most consistently considered were stage and age (6 of 7 models); gender and lymphocyte count (4 of 7 models); and serum albumin, B symptoms, hemoglobin (or hematocrit), and bulk of mediastinum (3 of 7 models). Conversely, erythrocyte sedimentation rate, serum lactate dehydrogenase, inguinal involvement, number and distribution of involved sites, and white blood cell count were present in no more than one index. The reason why this latter group of characteristics has remained scattered is probably either that they have been scrutinized by too few investigators or that their significance has been absorbed by other parameters in multivariable analyses. On the whole, a convergence of at least 2 of 7 studies on 10 prognostic factors is not a disappointing result when considering the biologic variability of HD as well as the methodologic heterogeneity already outlined. A similar trend of convergence for some prognostic factors was recorded by Carde[10,25] in a review of chemotherapy trials.

Third (and this is the major problem, as noted previously by Hasenclever and Diehl[9,13]), all of the prognostic factors in patients with HD, despite their highly statistically significant correlation with the time dependent variable, show rather low predictive power. This would explain the difficulty in identifying the groups of patients with a very high or a very low failure risk. Parameters that are found to fit the data best in the original set can prove to be less adequate and less reliable in a different set. With reference to the IPFP categories, patients who presented with four or more unfavorable factors, who were 19% of the population of patients in the original study[9] and showed 47% TTF, actually corresponded to only 10% of the patients with a 43% risk of failure in the current study sample or to 16% of the test sample with a failure risk of 51%. This is the main drawback with prognostic factors in patients with HD: poor predictive ability despite very significant correlation with clinical behavior. Probably, this fact can be related to the complexity of the disease, with the variable coexistence of both neoplastic and inflammatory components that affects histology, biology, and clinical manifestations and explains the high number of aspecific factors with which we still are dealing in clinical practice.

A further source of low prediction ability may be the categorization often made, within a few discrete ranges, of the continuously distributed values of clinical and laboratory variables. When a variable is categorized, simplicity always takes the upper hand, although a certain amount of information often is lost. Moreover, values of laboratory parameters close to a given cut-off may have their positive or negative impact on prognosis strongly magnified, because they fall just a bit over or under that cut-off value. With this mechanism, multiple combinations of categorized levels of different variables can lose further potential discrimination ability through the rise of a sort of prognostic background noise. This happens especially

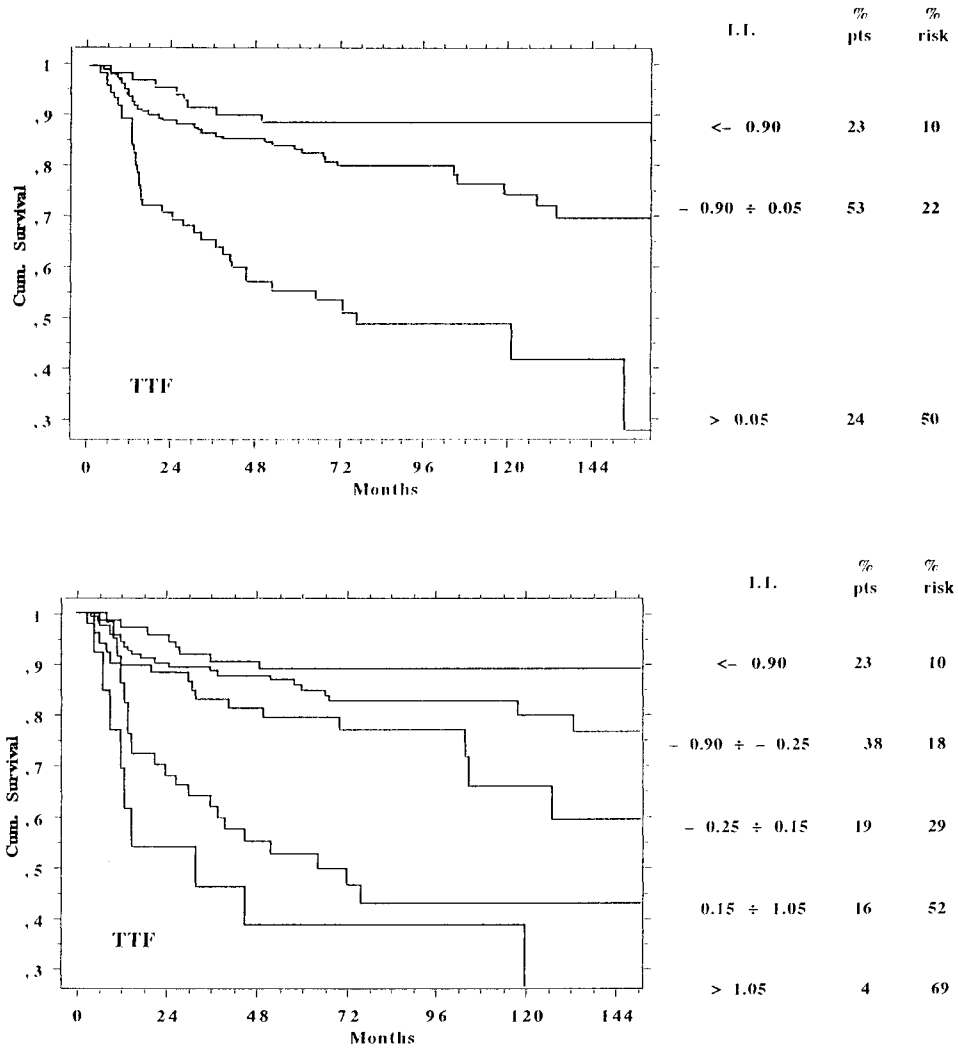

FIGURE 3. Examples of selecting patients with different prognoses by choosing different cut-off values for the integrated index (I.I.) value. The higher the I.I. the lower the percentage of patients (pts) with an I.I. greater than that value and the more severe the prognosis, as shown by time to treatment failure (TTF) curves and percent of cumulative (Cum.) risk.

when the cut-off value is not too far from the normal level, and the probability of finding values strictly around the cut-off value is high. This may be one of the principal reasons for the superiority of the IDHD model[7] over the PV-MO model,[3] as demonstrated previously by Bettini at al.[34] Conversely, the choice of clearly abnormal cut-off values for several variables can improve the ability of predicting a truly severe prognosis while at the same time reducing the number of patients for whom such a prediction is possible. This may be the case for the MSK index,[5] in which some factors were chosen with severely abnormal cut-off values, such as that for serum lactic dehydrogenase ($>$ 100% above the normal level), that for the definition of mediastinal bulk ($\geq$ 0.45 of the maximal chest width instead of 0.33, as currently used), and that for hematocrit (16–17% lower than normal). This may explain why, in the original MSK series, Straus et al.[5] found only 13 of 161 patients (8%) who presented with three or more unfavorable factors, but the prognosis of those 13 patients was extremely severe (20% OS at 5 years). In the current series, there were 29 of 315 patients (9%) with the same characteristics in the study sample and 24 of 201 patients (12%) in the test sample with an identical corresponding 0.63 risk of unfavorable events.

Differences in normal range limits and/or shifting from normal distribution for a quantitative prognostic factor are insidious sources of a slight reduction in predictive ability when different patient populations are studied. However, when a laboratory parameter is proposed and used as the percentile of frequency distribution, standardization is optimal and absolute. In this regard, serum albumin distribution percentiles requested for the IDHD index statistically are the most correct way to standardize albumin to the range of its observed values, contrary to what was argued by Fermé et al.[11] A similar problem was encountered by Wagstaff et al.[4] when dealing with the peripheral lymphocyte count. This showed rather different distribu-

tions among the two populations from two cooperating hospitals, and the investigators decided to work out the problem by choosing a cut-off level at which the prognostic effect was similar in the two centers. This choice represents a potential difficulty for any other center with a different best cut-off level for the correlation between lymphocyte count and prognosis.

On the whole, to date, the reality of prognostic investigation in patients with advanced HD is that only a part—and not a very great part–of the outcome variability can be predicted, and that a number of factors prove to be undoubtedly correlated with disease prognosis, but they show low predictive power for a number of reasons other than the intrinsic biologic and clinical variability of the disease (differences in time dependent variables chosen, in the analysis criteria adopted, in the type of clinical and biologic parameters scrutinized, in the statistical techniques used to handle the variables, etc.).

So long as therapy for patients with HD relies on pharmacologic and/or physical agents with dose limiting early and late toxicity, thereby continuing to justify the research on prognostic determinants,[10,25] the present difficulty in improving our predictive ability can be solved only by finding new prognostic factors that can explain a greater proportion of prognostic variability. Parameters such as tumor burden, serum CD30 level, or serum concentration of certain cytokines hopefully may be able to reflect disease biology or host-disease correlations more properly and closely than the prognostic factors in present use. However, because the selection, validation, and integration of possible new factors (if any emerge) probably will take many years, an effort to optimize the use of the prognostic information attainable from the factors and indexes currently available seems to be reasonable. The integration of three distinct models comes from the observation that not only do they prove to fit the data on patients with advanced HD better than any of the others, but they actually use 14 of the 15 covariates taken into account by all of the prognostic systems studied; moreover, they derive from studies that involved 10,349 of the total 12,415 patients examined in the original series from which the seven models were drawn. These may be some of the reasons why the prognostic effectiveness of the integrated index did not reduce in the test sample population, which generally happens for any index when working outside the set of patients originally analyzed. Furthermore, it is a flexible tool, because it allows the physician to choose any preferred combination of a given clinical risk level with the proportion of patients who can be expected with that risk, so it can be employed both for planning therapeutic strategies and for comparing results from clinical trials. Certainly, it reflects a kind of prognostic "syncretism" to which we are compelled—to try to get the maximum from the best available—as we await some new, reliable prognostic achievement to furnish a substantial step forward. Thus, the integrated index is not a new, "magical formula" but a selection of the best formulas already existing to provide the most reliable guide to currently available therapeutic options and the most accurate basis for the evaluation of future improvements.

ADDITIONAL INFORMATION

A program implemented on an electronic sheet avoids complex calculations when all of the necessary clinical data are available from a series of patients with advanced HD. The program can be downloaded from the web site http://www.unimo.it/gisl/default.htm. In practice, the clinical staging parameters required are the following: gender (male or female), age (number of years), histology (four histologic subtypes: LP, NS, MC, and LD), clinical stage (II–IV), systemic symptoms (absent [A] or present [B]), number and distribution of the involved anatomic lymph node areas (IAD: three or fewer areas above the diaphragm or otherwise), serum albumin concentration (g/dL), mean and standard deviation of the serum albumin concentration of a series of patients evaluated in the same center (g/dL), serum lactic dehydrogenase (LDH; U/L), hematocrit level (Ht; %), hemoglobin concentration (g/dL), white blood cell count (WBC; no. $\times 10^9$/L), peripheral blood lymphocyte count (Ly; no. $\times 10^9$/L and % of the WBC count), inguinal lymph node involvement (absent or present), bulky mediastinum involvement (absent or present), and bone marrow involvement (absent or present). The obtained integrated index has to be reported on the nomogram curves (see Fig. 2): the first curve provides a better estimate of the good-risk tail of the patient population, and the second curve provides a better estimate the poor-risk tail. On the chosen curve, the integrated index value will identify the corresponding combination of the expected risk and the percent of patients exposed to that risk. The risk involves the following events: disease progression during treatment, partial or null response after therapy, disease recurrence after variable time from achievement of complete remission, or death for the disease at any time.

REFERENCES

1. Urba WJ, Longo DL. Hodgkin's disease. *N Engl J Med* 1992; 326:678–87.
2. Canellos GP. The second chance for advanced Hodgkin's disease. *J Clin Oncol* 1992;10:175–7.

3. Gobbi PG, Cavalli C, Federico M, Bertoloni D, Di Prisco UA, Rossi A, et al. Hodgkin's disease prognosis: a directly predictive equation. *Lancet* 1988;i:675–9.
4. Wagstaff J, Gregory WM, Swindell R, Crowther D, Lister TA. Prognostic factors for survival in Stage III B and IV Hodgkin's disease: a multivariate analysis comparing two specialist treatment centres. *Br J Cancer* 1988;58:487–92.
5. Straus DJ, Gaynor JJ, Myers J, Merke DP, Caravelli J, Chapman D, et al. Prognostic factors among 185 adults with newly diagnosed advanced Hodgkin's disease treated with alternating potentially noncross-resistant chemotherapy and intermediate-dose radiation therapy. *J Clin Oncol* 1990; 7:1173–86.
6. Proctor SJ, Taylor P, Donnan P, Boys R, Lennard A, Prescott RJ, with members of the Scotland and Newcastle Lymphoma Group (SNLG) Therapy Working Party. A numerical prognostic index for clinical use in identification of poor-risk patients with Hodgkin'disease at diagnosis. *Eur J Cancer* 1991;27:624–9.
7. Gobbi PG, Comelli M, Grignani GE, Pieresca C, Bertoloni D, Ascari E. Estimate of expected survival at diagnosis in Hodgkin's disease: a means of weighting prognostic factors and a tool for treatment choice and clinical research. A report from the International Database on Hodgkin's Disease (IDHD). *Haematologica* 1994;79:241–55.
8. Lee SM, Radford JA, Ryder WDJ, Collins CD, Deakin DP, Crowther D. Prognostic factors for disease progression in advanced Hodgkin's disease: an analysis of patients aged under 60 years showing no progression in the first 6 months after starting primary chemotherapy. *Br J Cancer* 1997;75: 110–5.
9. Hasenclever D, Diehl V, for the International Prognostic Factors Project on Advanced Hodgkin's Disease. A prognostic score to predict tumor control in advanced Hodgkin's disease. *N Engl J Med* 1998;339:1506–14.
10. Carde P. Who are the high-risk patients with Hodgkin's disease? *Leukemia* 1996;10(Suppl 2):S62–7.
11. Fermé C, Bastion Y, Brice P, Lederlin P, Divine M, Gabarre J, et al. Prognosis of patients with advanced Hodgkin's disease. Evaluation of four prognostic models using 344 patients included in the Group d'Etudes des Lymphomes de l'Adulte Study. *Cancer* 1998;80:1124–33.
12. Fisher RI. Report of Workshop 5: early high-dose chemotherapy. *Ann Oncol* 1996;7(Suppl 4):131–3.
13. Haisenclever D, Diehl V. The International Prognostic Factors Project on advanced Hodgkin's disease: a prognostic index, but no very high risk group [abstract]. *Ann Oncol* 1996;7(Suppl. 3):21(65).
14. Hoppe RT. Development of effective salvage treatment programs for Hodgkin's disease: an ongoing challenge. *Blood* 1991;77:2093–5.
15. Gobbi PG, Pieresca C, Federico M, Di Renzo N, Narni F, Iannitto E, et al. MOPP/EBV/CAD hybrid chemotherapy with or without limited radiotherapy in advanced or unfavorably presenting Hodgkin's disease: a report from the Italian Lymphoma Study Group. *J Clin Oncol* 1993;11:712–9.
16. Gobbi PG, Pieresca C, Ghirardelli ML, Di Renzo N, Federico M, Merli F, et al. Long-term results from MOPPEBVCAD chemotherapy with optional limited radiotherapy in advanced Hodgkin's disease. *Blood* 1998;91:2704–12.
17. Mazza P, Bocchia M, Zinzani PL, Fiacchini M, Gherlinzoni F, Bendandi M, et al. Hodgkin's disease: summary of twenty years' experience. *Haematologica* 1992;77:487–93.
18. Lambertenghi-Deliliers G, Baldini L, Radaelli F, Federico M, Polli E. Efficacia terapeutica della polichemioterapia CcVPP nel morbo di Hodgkin. In: Mauri C, Silingardi V, Federico M, editors. La malattia di Hodgkin. Recenti acquisizioni e problemi aperti. Modena: Mucchi, 1987:201–7.
19. Gobbi PG, Cavalli C, Gendarini A, Crema A, Ricevuti G, Federico M, et al. Re-evaluation of prognostic significance of symptoms in Hodgkin's disease. *Cancer* 1985;56:2874–80.
20. Tesch H, Sieber M, Diehl V. Treatment of intermediate stages. In: Diehl V, editor. Hodgkin's disease. Baillere's clinics in haematology. London: WB Saunders Co., 1996:543–51.
21. Hoppe RT, Coleman CN, Cox RS, Rosenberg SA, Kaplan HS. The management of Stage I–II Hodgkin's disease with irradiation alone or combined modality therapy: the Stanford experience. *Blood* 1982;59:455–6.
22. Longo DL, Russo A, Duffey PL, Hubbard SM, Glatstein E, Hill JB, et al. Treatment of advanced-stage massive mediastinal Hodgkin's disease: the case for combined modality treatment. *J Clin Oncol* 1991;9:227–35.
23. Gause BL, Longo DL. Treatment of relapsed Hodgkin's disease. In: Diehl V, editor. Hodgkin's disease. Baillere's clinics in haematology. London: WB Saunders Co., 1996:559–72.
24. Lister TA, Crowther D, Sutcliffe SB, Glatestein E, Canellos GP, Young RC, et al. Report of the committee convened to discuss the evaluation and staging of patients with Hodgkin's disease: Cotswolds Meeting. *J Clin Oncol* 1989;7: 1630–6.
25. Carde P. Should poor risk patients with Hodgkin's disease be sorted out for intensive treatments? *Leukemia Lymphoma* 1995;15:31–40.
26. Longo DL, Duffey PL, Young RC, Hubbard SM, Ihde DC, Glatstein E, et al. Conventional-dose salvage combination chemotherapy in patients relapsing with Hodgkin's disease after combination chemotherapy: the low probability for cure. *J Clin Oncol* 1992;10:210–8.
27. Dixon DO, McLauglin P, Hagemeister FB, et al. Reporting outcomes in Hodgkin's disease and lymphoma. *J Clin Oncol* 1987;5:1670–2.
28. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc* 1958;53:457–81.
29. Cox D. Regression model and life-tables (with discussion). *J R Stat Soc* 1972;34:187–220.
30. Kaplan HS. Hodgkin's disease. 2nd ed. Cambridge: Harvard University Press, 1980.
31. Carde P, MacKintosh FR, Rosenberg SA. A dose and time response analysis of the treatment of Hodgkin's disease with MOPP chemotherapy. *J Clin Oncol* 1983;1:146–53.
32. Carella AM, Prencipe E, Pungolino E, Lerma E, Frassoni F, Rossi E, et al. Twelve years experience with high-dose therapy and autologous stem cell transplantation for high-risk Hodgkin's disease patients in first remission after MOPP/ABVD chemotherapy. *Leukemia Lymphoma* 1996;21:63–70.
33. Klimo P. Evolution of the concept of prognostication in therapy of Hodgkin's lymphoma. *Leukemia Lymphoma* 1995;15:15–8.
34. Bettini R, Tonolini M, Maccianti E. Verification and comparison of two different predictive equations in Hodgkin's disease. *Haematologica* 1997;82:324–7.