New prognostic scoring system for primary myelofibrosis based on a study of the International Working Group for Myelofibrosis Research and Treatment

Francisco Cervantes,1 Brigitte Dupriez,2 Arturo Pereira,1 Francesco Passamonti,3 John T. Reilly,4 Enrica Morra,5 Alessandro M. Vannucci,6 Ruben A. Mesa,7 Jean-Loup Demory,2 Giovanni Barosi,6 Elisa Rumi,3 and Ayalew Tefferi7

1Hematology Department, Hospital Clinic, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), University of Barcelona, Barcelona, Spain; 2Centre Hospitalier, Lens and Lille, France; 3Hematology Division, Fondazione Istituto di Ricovero e Cura a Carattere Scientifico (IRCCS) Policlinico S Matteo, University of Pavia, Pavia, Italy; 4Royal Hallamshire Hospital, Sheffield, United Kingdom; 5Università Milano-Niguarda, Milan, Italy; 6University of Florence, Florence, Italy; 7Division of Hematology, Mayo Clinic College of Medicine, Rochester, MN; and 8Unit of Clinical Epidemiology, Center for the Study of Myelofibrosis, Fondazione IRCCS Policlinico S Matteo, Pavia, Italy

Therapeutic decision-making in primary myelofibrosis (PMF) is becoming more challenging because of the increasing use of allogeneic stem cell transplantation and new investigational drugs. To enhance this process by developing a highly discriminative prognostic system, 1054 patients consecutively diagnosed with PMF at 7 centers were studied. Overall median survival was 69 months (95% confidence interval [CI]: 61-76). Multivariate analysis of parameters obtained at disease diagnosis identified age greater than 65 years, presence of constitutional symptoms, hemoglobin level less than 10 g/dL, leukocyte count greater than 25 × 10⁹/L, and circulating blast cells 1% or greater as predictors of shortened survival. Based on the presence of 0 (low risk), 1 (intermediate risk-1), 2 (intermediate risk-2) or greater than or equal to 3 (high risk) of these variables, 4 risk groups with no overlapping in their survival curves were delineated; respective median survivals were 135, 95, 48, and 27 months (P < .001). Compared with prior prognostic models, the new risk stratification system displayed higher predictive accuracy, replicability, and discriminant power. In 409 patients with assessable metaphases, cytogenetic abnormalities were associated with shorter survival, but their independent contribution to prognosis was restricted to patients in the intermediate-risk groups. JAK2V617F did not cluster with a specific risk group or affect survival. (Blood. 2009;113: 2895-2901)

Introduction

Primary myelofibrosis (PMF)1 is classified as a chronic myeloproliferative disorder and characterized by variable degrees of cytopenia(s) and/or cytosis, a leukoerythroblastic blood picture, bone marrow fibrosis, and extramedullary hematopoiesis often resulting in hepatosplenomegaly.2 From a pathogenesis standpoint, the disease features clonal proliferation involving pluripotent hematopoietic stem cells,3,4 and clonal cell–derived cytokines are implicated for some of the disease aspects such as bone marrow fibrosis and extramedullary hematopoiesis.2 Most recently, JAK25-7 and MPL8-10 mutations were described in approximately 50% and 10% of patients with PMF, respectively. However, the precise pathogenetic contribution of these mutations is currently not well defined.

PMF usually affects subjects with advanced age,11 but young people are not necessarily spared.12 Reported median survivals are variable and in the range of 4-7 years.13,14 Previous studies have identified several adverse prognostic factors for survival, including advanced age,15-19 marked anemia,13-22 leukocytosis or leukopenia,14,16,18,22 abnormal karyotype,18,23-25 constitutional symptoms,13,14,17,22 and presence of circulating blasts.13,14,22 Based on some of these variables, several prognostic scoring systems have been proposed.12,13,18,19,22,26 More recently, the prognostic value of blood CD34+ cell count27,28 and JAK2 mutational status29-31 has also been evaluated.

Current drug therapy for PMF has not been shown to influence survival and is often used for palliative purposes only.32 To get out of this therapeutic deadlock, there is growing use of allogeneic stem cell transplantation (allo-SCT)33-38 and, more recently, anti-JAK2–targeted therapy.32 Patient selection for these and other therapeutic approaches in PMF is often challenging and is the main reason to undertake the current large multicenter study, to accurately identify prognostic factors that would facilitate therapeutic decision making for the individual patient.

Methods

Patients and diagnostic criteria

After approval from the Institutional Review Board of each participating study center, the databases of the 7 participating institutions were screened, and a total of 1131 consecutive patients diagnosed with PMF during the period of January 1980 to April 2007 were analyzed. By definition, cases of post-polycythemia vera (post-PV) or post–essential thrombocythemia (post-ET) myelofibrosis4 were not considered. After a systematic individual case review, 19 patients were excluded due to the following reasons: 3 patients met criteria for blastic transformation at presentation (blast cells in bone marrow or blood ≥20%), 1 had extreme leukocytosis (>150 × 10⁹/L), 9 met criteria for chronic myelomonocytic leukemia, and 6 had hemoglobin levels that would qualify for a diagnosis of PV. In addition, 58 more patients were excluded by applying the diagnostic criteria for PMF recently established by the World Health Organization (WHO) classification system,39 because diagnosis in these excluded cases had been based on the...
presence of isolate thrombocytosis (>450 × 10^9/L) with grade 1 bone marrow fibrosis that was not associated with palpable splenomegaly, anemia, leukoerythroblastosis, or increased serum lactate dehydrogenase (LDH) levels. In the end, therefore, 1054 patients were submitted for analysis of prognostic factors.

Due to the long study period, diagnosis of PMF was made according to the criteria accepted at the time when the patient was diagnosed. In all cases, presence of an increased reticulin and/or collagen bone marrow content without any apparent cause (such as chronic myeloid leukemia, PV, myelodysplasia, lymphoproliferative disorders, scleroderma, primary pulmonary hypertension, or others) was required, in addition to the presence of features typical of the disease, including palpable splenomegaly, leukoerythroblastosis, or histologic demonstration of myeloid metaplasia. Cases of the so-called “prefibrotic” form of PMF, characterized by lack of marrow fibrosis with highly dysplastic megakaryocytes, usually accompanied by thrombocytosis, but without anemia, splenomegaly, or leukoerythroblastosis, were not considered, because most patients were diagnosed before this histologic variant of PMF was recognized by the WHO classification and to minimize the inadvertent inclusion of patients with ET.

### Treatment

Disease management was variable and usually based on the disease characteristics in every individual patient. It included a wait-and-see approach until disease progression in asymptomatic patients, single-agent oral chemotherapy (mainly hydroxyurea, but also busulfan, 6-mercaptopurine, pipobroman, and thioguanine), androgens, erythropoiesis-stimulating agents, prednisone, interferon-α, anagrelide, immunomodulatory agents such as thalidomide and lenalidomide and, in few instances, intravenous cytotoxic agents such as radiophosphorus and cladribine. A total of 111 patients underwent splenectomy during the study period, 5 received allo-SCT with a standard (n = 4) or a reduced-intensity conditioning (RIC) regimen (n = 1), and 2 autologous SCT.

### Data evaluated

The initial variables selected for prognostic assessment were those previously shown to be of prognostic value in PMF; those clinically meaningful, and possible confounders (namely, diagnostic period and series of origin), provided they were available in the majority of patients. With the above premises, the following characteristics were analyzed for prognostic significance: diagnostic period (before and after 1995), institution of origin, patient’s sex, age greater than 65 years, presence of constitutional symptoms (weight loss > 10% of the baseline value in the year preceding PMF diagnosis and/or unexplained fever or excessive sweats persisting for more than 1 month), hemoglobin (Hb) less than 10 g/dL, leukocyte count (considered at the cutoff levels of 4 × 10^9/L, 20 × 10^9/L, 25 × 10^9/L, and 30 × 10^9/L), presence (> 1%) of circulating blasts in peripheral blood, and platelet count less than 100 × 10^9/L. These variables were available in 96.6% to 100% of patients. At a further step, other variables available in a lower proportion of patients were also analyzed, including monocyte counts greater than 1 × 10^9/L (counts available in 65.4% of cases), presence or not of abnormalities in a karyotype obtained from marrow or unstimulated blood (as assessable in 409 patients), JAK2 mutational status (345 patients), and blood CD34+ cell count (150 cases).

### Statistical methods

The major outcome was survival from diagnosis, and it was estimated using the Kaplan-Meier method. The effect of the potential prognostic factors on survival was evaluated by the Cox proportional hazards (PH) regression. In every Cox model, the PH assumption was checked by graphical methods and by the Grambsch-Therneau test.

The final prognostic model was identified through a stepwise selection process based on a z test of the regression coefficients. Initially, all potential risk factors and confounders were included in the multivariate Cox model. At each step, variables with a P value for the z test greater than the cutoff were excluded from the model, and the remaining ones were tested again for their independent association with survival until no more variables met the criteria for exclusion. To safeguard against associations occurring by chance due to multiple simultaneous tests, the cutoff values for the z test were Bonferroni-adjusted by dividing 0.05 by the number of covariates in the model at each step.

The prognostic scoring system was evaluated by calculating its discriminating power, compared with that of currently used PMF prognostic systems, and its positive predictive accuracy for actual survival longer (or shorter) than definite time periods from diagnosis. The discriminating power was measured by the Harrell’s C concordance index, which represents the proportion of all possible pairs of patients in which the ordering of the risk of death, as predicted by the model, agrees with the observed outcome, after excluding tied observations. Values can range from 0 to 1, with values close to 1 indicating that the scoring system almost perfectly discriminates between patients with different risk of death, while those close to 0.5 indicate that the model’s discriminating power is not better than chance alone. Bias-corrected 95% confidence intervals (CIs) for the estimated Harrell index were calculated using resampling with 500 replicates.

The positive predictive accuracy of a prognostic subgroup for actual survival is the proportion of patients in this subgroup who survive longer (or shorter) than a given time cutoff. It was calculated after excluding patients censored before the cutoff, whose actual survival cannot yet be known for certain, as previously described.

The replicability of the prognostic scoring system was tested by bootstrap resampling. One thousand samples, the same size as the original series, were built through random extraction with reposition, so that in each sample, a given patient might either not be represented at all or represented once, twice, or more times. The parameters assessed by resampling were the 95% CI for the hazard ratio of the prognostic factors identified at the Cox regression model and the 95% CI for the 75th, 50th, and 25th percentiles of survival time for each prognostic subgroup. Resampling allows verifying that the prognostic factors identified at the Cox model and the derived prognostic subgroups were not critically dependent on the particular composition of the study series.

The relative survival is the ratio of the mortality observed in the series to the yearly age- and sex-adjusted mortality of the general population for the country of origin and background life span after diagnosis. It permits to evaluate the intrinsic prognostic value of variables like age and sex after excluding their demographic-driven effect on mortality and to compare the survival of patients diagnosed over distant time periods or in countries with a different background life expectancy. Relative survival analysis allows to identify disease-specific prognostic factors even whether the ultimate cause of death can be attributed or not to the disease under study. Relative survival curves were computed using the method described by Hakulinen, and the relative-survival values used in the multivariate analysis were calculated by the Ederer II method. The independent association of the potential prognostic factors with relative survival was evaluated by multivariate Poisson regression according to the methods described by Dickman et al.

Country-specific yearly age- and sex-adjusted mortality rates from the year of diagnosis were obtained from the Human Mortality Database (http://www.mortality.org).

Comparisons between variables were done using the Chi-square test for variables expressed as proportions and the Mann-Whitney test for ordered or continuous variables. The log-rank test was used to compare Kaplan-Meier survival curves. All tests were 2-sided, and P values less than .05 were considered significant. All analyses were conducted using the STATA software (http://www.stata.com). For relative survival analysis, the STATA routines developed by Dickman (Karolinska Institutet, Stockholm, Sweden; available at http://www.pauldickman.com) were used.

### Results

**Patients’ features at presentation**

Table 1 summarizes the main clinical and laboratory features of the 1054 patients at diagnosis. As can be seen, the median age was 64 years; 178 patients (16.9%) were younger than 50 years, and
Prognostic factors

In the first step of the stepwise Cox model, 3 variables were excluded: series of origin, diagnostic period, and white blood cell (WBC) count less than $4 \times 10^9/L$. The variable sex was removed in the second step. No additional variables were excluded in the third step. No additional variables were excluded in the third step. No additional variables were excluded in the third step.

Survival and causes of death

At the time of analysis, 517 patients (49%) had died. Figure 1 shows the actuarial survival curve of the series. Median survival was 69 months (95% CI: 61-76). Among patients in whom the final cause of death was known, transformation to acute leukemia was the most frequent cause (86 patients), followed by PMF progression without acute transformation (50 cases), thrombosis and cardiovascular complications (37 cases), infection (n = 29) or bleeding (n = 14) out of the setting of acute transformation, portal hypertension (n = 12), and other causes (n = 48, including 12 cases of second neoplasias). Two patients died from complications of transplantation.

Prognostic score

Because there were no marked differences in the hazard ratios of the 5 prognostic variables, for the sake of simplicity, 1 point was assigned to each one of them. As a result, the patients’ score could range from a minimum of 0 to a maximum of 5. Only 4 patients in the series had a score of 5. In addition, overlapping was observed in the 95% CI of the median survival of patients with scores 3 and 4, whereas no such overlapping was seen in patients with scores 0, 1, and 2. Therefore, patients with scores 3, 4, and 5 were pooled into a single group, and 4 prognostic groups were finally considered: low risk (no poor prognostic factor, including 22% of the patients; median survival 135 months); intermediate risk-1 (1 poor prognostic factor, 29% of the patients; median survival 95 months); intermediate risk-2 (2 prognostic factors, 28% of the patients; median survival 48 months); and high risk (3 or more prognostic factors, 21% of the patients; median survival 27 months; $P < .001$; Table 3 and Figure 2).

Figure 1. Actuarial survival curve of the overall series.
Several variations of the model were tested. Thus, when 70 years was used as the cutoff for age, a decrease in the discriminating power of the model was observed. This was even more pronounced when the variables age and constitutional symptoms were removed from the model. When the prognostic value of the monocytosis was tested in the subgroup of 675 patients with the data, it did not increase the prognostic weight of the model. Patients without splenomegaly at diagnosis survived longer than the remainder, but the difference did not reach statistical significance, whereas the variable splenomegaly did not improve the PMF prognostic score. On the other hand, the possible effect of splenectomy in the patients’ evolution was not evaluated, as the study was designed to analyze the prognostic significance of presenting and not evolutive data.

The results of the validation of the PMF prognostic score are shown in Figures S2 and S3.

Analysis of the relative mortality

Figure 3 depicts the relative survival of the series compared with that of the general population. As can be seen, once the demographic effects of age, sex, country of origin, and year of diagnosis were excluded, a marked effect of the disease on the patients’ survival was observed. Indeed, mortality at 5 and 10 years from diagnosis was 40% and 60% greater, respectively, than the expected mortality in a general population with similar demographic characteristics. Beside, the 5 prognostic factors identified at the stepwise Cox regression model also proved to be the ones significantly influencing on relative survival (data not shown).

Other variables analyzed

In patients with available karyotype (n = 409), presence of cytogenetic abnormalities was associated with Hb less than 10 g/dL ($P < .001$) and showed a significant association with shorter survival even after adjustment for the prognostic score ($P = .01$). Of note, the presence of an abnormal karyotype contributed to the prognosis, but only in the 2 intermediate-risk groups and not in the high- and low-risk groups.

In patients assessed for $JAK_2$ status, a significant association was found between presence of the mutation and age greater than 65 years ($P < .002$). However, no association was observed between $JAK_2$ status and the prognostic score or the survival.

In patients with available blood CD34$^+$ cell count at diagnosis, this parameter correlated with blood blasts greater than or equal to 1% ($P = .004$), but not with prognostic score or survival. Although CD34$^+$ cell counts greater than $300 \times 10^9/L$ (n = 11) were associated with shorter survival, their prognostic value disappeared when adjusted by prognostic score.

Discussion

PMF is a heterogeneous disease in its presentation and evolution. Median survival is highly variable; a proportion of patients die shortly after diagnosis, whereas a few survive for 2 decades or longer. This fact has stimulated identification of prognostic factors and, as a result, several prognostic systems have been proposed. As PMF treatment is essentially palliative, allo-SCT is increasingly being used, and newer drugs are being tested in these patients. Although the mortality of RIC allo-SCT is lower than that of conventional allo-SCT, there are still some...
associated mortality and morbidity. Therefore, prognostic stratification of PMF patients is important to make treatment decisions. However, given the low frequency of the disease, to date, prognostic studies have been performed in series including a relatively small number of patients. In this sense, the present study, in which more than 1000 patients from 7 institutions were analyzed, represents the largest prognostic study ever performed in this disease.

Median survival of patients in the present series was 5.7 years. Patients diagnosed after 1995 survived slightly longer than those diagnosed before \( (P = .045) \), but the prognostic significance of date of diagnosis disappeared at multivariate analysis. As expected from previous observations, an initial hemoglobin level of less than 10 g/dL was the variable with the highest impact on survival.\(^{13,22} \) Presence of constitutional symptoms is also another previously well-established risk factor in PMF.\(^{13,14,17,22} \) It is to be noted that this parameter includes objective measures such as weight loss and fever and that the temptation to include it based on its subjective elements would significantly reduce the discriminating power of the current prognostic model. In the current study, leukocyte count greater than 25 \( \times 10^9/L \) performed better, as an adverse risk factor, compared with the previously described cutoff level of 30 \( \times 10^9/L. \)\(^{14,22} \) In addition, the new leukocyte threshold resulted in a higher number of informative patients in terms of patient stratification. In contrast, we could not confirm the poor prognostic influence of either low leukocyte\(^{14,26} \) or high monocyte\(^{26} \) counts. Presence of circulating blasts at presentation had an unfavorable prognostic influence, as previously shown.\(^{13,14,22} \) Finally, we were able to demonstrate that the adverse effect of advance age (age > 65 years) was not merely a demographic phenomenon but probably an indication of decreased tolerance to the disease and its complications by the elderly.

An abnormal karyotype was associated with shorter survival in the current study. However, we are mindful of the difficulty obtaining assessable metaphases in PMF because of a “dry tap” during bone marrow biopsy. The negative prognostic influence of cytogenetic abnormalities in PMF has previously been pointed out.\(^{18,23-25} \) In the series by Tefferi et al,\(^{25} \) such adverse influence was observed for cytogenetic abnormalities as a whole, but when chromosome changes were analyzed separately, the unfavorable influence was restricted to trisomy 8 and deletion of 12p, while deletions of 13q and 20q were not associated with shorter survival. In our study, the prognostic influence of specific karyotypic alterations could not be analyzed, because detailed information on the type of abnormality was not available from all contributing centers.

The prognostic impact of \( JAK2/V617F \) or its allele burden in PMF is currently being debated.\(^{29-31} \) In one retrospective study of 152 patients,\(^{29} \) shorter survival was associated with the presence of \( JAK2/V617F, \) whereas Barosi et al,\(^{30} \) in a prospective study of 174 patients, found a correlation between the mutation and evolution toward large splenomegaly, need for splenectomy, and frequency of leukemic transformation, but not survival. Similarly, in 117 patients from the Mayo Clinic,\(^{31} \) no prognostic value for the \( JAK2 \) mutation was noted. In the 345 patients from the present study in which \( JAK2/V617F \) mutational status was available, the presence of the mutation was not associated with either prognostic score or survival. Information on \( JAK2/V617F \) allele burden was available in a small proportion of study patients, and differences among study centers regarding assay methodology and cell types used to measure allele burden prevent valid prognostic analysis using the particular parameter in the current study.

PMF patients have an increased number of CD34\(^+ \) cells in peripheral blood.\(^{27} \) In one study,\(^{27} \) a correlation between circulating CD34\(^+ \) cell count and patient risk group was noted, with the higher the number of such cells the more unfavorable the risk group, but such prognostic correlation was not confirmed in the series of the Mayo Clinic,\(^{25} \) or in the current study.

Several prognostic score systems for PMF have been proposed.\(^{12,13,18,19,22,26} \) The most widely used is the “Lille score” reported by Dupriez et al,\(^{22} \) which features 3 prognostic categories based on hemoglobin level and leukocyte count, median survivals in low-, intermediate-, and high-risk groups being 93, 26, and 13 months, respectively. Subsequent studies, including the current one, showed that the Lille score does not clearly separate intermediate- and high-risk patient groups. More recently,\(^{26} \) the Mayo Clinic group tried to improve the Lille score by adding thrombocytopenia (\(< 100 \times 10^9/L \) and monocytosis (\( > 1 \times 10^9/L \)) as additional adverse risk factors. This resulted in better, but still suboptimal, separation of intermediate- and high-risk categories. Finally, the scoring system by Cervantes et al,\(^{12} \) applicable also to younger patients,\(^{25} \) is based on hemoglobin level and the presence or absence of constitutional symptoms and circulating blasts. However, the value of this system is limited by its ability to identify only 2 risk groups; identification of an intermediate-risk group is important, as shown in other hematologic diseases such as chronic myeloid leukemia\(^{30} \) or the myelodysplastic syndromes.\(^{32} \)

Based on 5 parameters readily available at time of diagnosis, the current prognostic model identifies 4 prognostic groups in PMF that are clearly different with regard to survival. Accordingly, low- and high-risk patient groups, including more than one-fifth of the study patients each, displayed respective median survivals of approximately 11 and 2 years, whereas median survivals of patients in the 2 intermediate-risk disease categories were 8 and 4 years. This new prognostic model had higher discriminating power than that seen with previous scoring systems and showed high replicability and predictive accuracy. Of note, the 2 intermediate-risk groups showed no overlapping in the survival curves and also a different influence on relative survival. Indeed, in patients in the intermediate-risk-1 group, the shortening in relative survival was observed only after 3 years of diagnosis, while in the intermediate-risk-2 group, such an effect was evident since time of presentation, supporting the consideration of 2 instead of 1 intermediate-risk groups. Also of note, cytogenetic findings had additional prognostic value in these intermediate-risk group patients only.

The new PMF prognostic system has considerable practical implications. For example, for low-risk patients, who have an expected median survival that exceeds 11 years, the risk treatment-related mortality and morbidity from allo-SCT or the possible toxicity of new investigational drug therapy might not be justified. On the other hand, it is reasonable to recommend investigational drug therapy for all other patients and allo-SCT for high- or intermediate-risk patients. Obviously, such recommendations as well as our proposed prognostic system are open to change based on additional new information.

Acknowledgments

This study was supported in part by grant RD06/0020/0004 from the Instituto de Salud Carlos III, Spanish Ministry of Health (Madrid, Spain). The International Working Group for Myelofibrosis Research and Treatment (IWG-MRT) meetings are sponsored by the M. D. Anderson Cancer Center Leukemia Department.
(Houston, TX) with additional funds from the Scott Richards Symposium proceedings (Houston, TX).

Authorship

Contribution: F.C. and A.T. designed research, interpreted results, and wrote the paper; A.P. performed the statistical analysis; and B.D., F.P., J.T.R., E.M., A.M.V., R.A.M., J.-L.D., G.B., and E.R. performed research and revised the manuscript.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Francisco Cervantes, Hematology Department, Hospital Clinic, Villarreal 170, 08036 Barcelona, Spain; e-mail: fcervan@clinic.ub.es; or Ayalew Tefferi, Division of Hematology, Mayo Clinic, 200 First St SW, Rochester, MN 55905; e-mail: tefferi.ayalew@mayo.edu

References


