bjh research paper

# Risk models predicting survival after reduced-intensity transplantation for myelofibrosis

**Haefaa Alchalby,[1] Dinah-Rohina Yunus,[1] Tatjana Zabelina,[1] Guido Kobbe,[2] Ernst Holler,[3] Martin Bornhäuser,[4] Rainer Schwerdtfeger,[5] Wolfgang Bethge,[6] Hans Michael Kvasnicka,[7] Guntram Büsche,[8] Francis Ayuk,[1] Ulrike Bacher,[1] Axel R. Zander[1] and Nicolaus Kröger[1]**

*[1]Department of Stem Cell Transplantation, University Medical Centre Hamburg-Eppendorf, Hamburg, [2]Department of Haematology and Oncology, University Hospital Düsseldorf, Düsseldorf, [3]Department of Haematology and Oncology, University Hospital Regensburg, Regensburg, [4]Department of Haematology and Oncology, University Hospital Dresden, Dresden, [5]Department of Stem Cell Transplantation, DKD Clinic, Wiesbaden, [6]Department of Haematology and Oncology, University Hospital Tübingen, [7]Institute of Pathology, University of Frankfurt, Frankfurt, and [8]Institute of Pathology, Hannover Medical School, Hannover, Germany*

Received 26 August 2011, accepted for publication 11 October 2011

Correspondence: Professor Dr. med. Nicolaus Kröger, Department of Stem Cell Transplantation, University Hospital Hamburg-Eppendorf, Martinistrasse 52, D-20246 Hamburg, Germany. E-mail: nkroeger@uke.uni-hamburg.de

## Summary

To define a prognostic model for predicting outcome of reduced-intensity allogeneic stem cell transplantation (RIC-ASCT) for myelofibrosis we evaluated 150 homogenously treated patients and developed a new risk score for overall survival (OS). In a multivariate Cox model for OS, only *JAK2* V617F wild-type, age ≥57 years and constitutional symptoms were independently predictive for OS (Hazard Ratio: 2·02; 2·43 and 2·80 respectively). Depending on the presence of one, two or all of these factors, HR of death was 3·08; 4·70 and 16·61 respectively ($P < 0·001$). This score was compared to the Lille, Cervantes, International Prognostic Scoring System (IPSS), dynamic IPSS (DIPSS) and modified European Blood and Marrow Transplantation Group (EBMT) scores. Lille score correlated significantly with OS but discriminated poorly between the intermediate and high-risk groups (5-year OS 56% and 51% respectively). IPSS and DIPSS correlated significantly with OS but differences between intermediate-1 and intermediate-2 groups were not significant (5-year OS 78% vs. 78% and 70%, 60% respectively). Modified EBMT and Cervantes models did not predict OS post-ASCT. In conclusion, a simple model which includes: age, *JAK2* V617F-status and constitutional symptoms can clearly separate distinct risk groups and can be used in addition to the Lille model to predict OS after RIC-ASCT for myelofibrosis.

**Keywords:** risk model, allogeneic stem cell transplantation, myelofibrosis.

Primary myelofibrosis (PMF) is a myeloproliferative neoplasm (MPN) that is characterized by a heterogenous clinical picture and a reduced life expectancy (Cervantes *et al*, 2008). Due to the highly variable natural history of the disease, treatment decisions are often challenging. Current drug therapy has failed to improve survival, and allogeneic stem cell transplantation (ASCT) is the only potentially curative option available for PMF patients, which however carries a substantial treatment-related mortality (TRM) (Deeg *et al*, 2003; Bacigalupo *et al*, 2010; Ballen *et al*, 2010). ASCT after reduced-intensity conditioning (RIC-ASCT) resulted in an improved TRM but which patients should be offered ASCT and the optimal time point for referral to ASCT is still debated (Kroger *et al*, 2009). With the availability of curative therapy, efficient tools for prediction of disease evolution became urgent and led to the development of several scoring systems such as the Lille score (Dupriez *et al*, 1996), Cervantes score (Cervantes *et al*, 1998), International Prognostic scoring system (IPSS) (Cervantes *et al*, 2009), dynamic IPSS (DIPSS) (Passamonti *et al*, 2010). These models were derived from patient groups treated mostly by a symptom-oriented approach. The European Blood and Marrow Transplantation Group (EBMT) score is a



First published online 27 January 2012
doi:10.1111/j.1365-2141.2011.09009.x

transplantation-specific stratification model that was introduced to apply efficiently in all haematological malignancies but has not been yet evaluated for patients with myelofibrosis (Gratwohl *et al*, 2009). Which scoring model is the most helpful for therapy decision and advising patients has not yet been systematically evaluated. In the last decade, several studies were published regarding ASCT for myelofibrosis, many of which tried to identify independent risk factors for negative outcome and two of them presented predictive scores for survival (Table I). Those efforts were unfortunately hampered by the small and heterogeneous study groups in both studies.

In the current study we propose a new scoring system to predict overall survival (OS) of patients with myelofibrosis undergoing RIC-ASCT using a rather homogenous cohort. Furthermore, we compared this scoring system to some other readily available and commonly used stratification systems for this disease.

## Materials and methods

### *Patients*

Data from 173 patients with histologically proven primary or post essential thrombocythaemia/polycythaemia vera (ET/PV) myelofibrosis who underwent RIC-ASCT between December 1999 and March 2010 were analysed. 122 patients were treated within two previously published clinical trials (Kroger *et al*, 2005, 2009). The remaining 51 patients were treated in one centre (Hamburg, Germany) according to the same treatment protocol but after closure of the previous studies. Approval was obtained from the local ethic committee for these studies and informed consent was provided according to the Declaration of Helsinki. Patient characteristics are summarized in Table II. In an attempt to define a risk model based upon a transplantation cohort we evaluated a subgroup of 150 patients from whom a near to complete data set was available, applying a univariate Cox analysis and then a multivariate Cox regression model to define independent risk factors for OS.

The performance of several readily available scoring systems, summarized in Table III, was also explored using the whole cohort. For the EBMT-Gratwohl score, disease stage was expressed as follows: early, when no circulating blasts are detectable; intermediate, when 1–20% blasts are detectable; and late, when peripheral blasts exceed 20%. The cut-off of time interval from PMF (or fibrotic transformation from PV or ET) to SCT was derived from the median corresponding value of our cohort because no myelofibrosis-specific definitions were available in the original score. As the cut-off age used in the IPSS system (65 years) seemed to be too old for a transplantation cohort, we reclassified this threshold by using a cut-off derived from the median age of the current group (57 years) as an age-adapted IPSS.

### *Statistical analysis*

The endpoint for this evaluation was OS. Differences in the distribution of categorical variables were compared using chi-square test and continuous variables were compared using

**Table I.** Independent risk factors for OS identified by several clinical trials which focused on ASCT for myelofibrosis.

| Trial | Conditioning used | Trial type | Independent risk factors for OS |
|---|---|---|---|
| Guardiola *et al* (1999) | Myeloablative | Retrospective | Hb ≤100 g/l<br>Fibrosis grade 3 |
| Deeg *et al* (2003)* | Myeloablative | Retrospective | Lille score<br>Degree of marrow fibrosis<br>Low platelet count |
| Kerbauy *et al* (2007) | Myeloablative and reduced-intensity | Retrospective | No Targeted busulfan and cyclophosphamide conditioning<br>Low platelet count<br>Advanced age<br>Increased Sorror morbidity score |
| Kroger *et al* (2009) | Reduced-intensity | Prospective | Age >55 years<br>Mismatched donor |
| Bacigalupo *et al* (2010)* | Reduced-intensity | Retrospective | Splenomegaly >22 cm<br>Blood transfusions >20 unit<br>Other than HLA-identical sibling donor |
| Alchalby *et al* (2010) | Reduced-intensity | Retrospective | Advanced age<br>Mismatched donor<br>No *JAK2* V617F |
| Robin *et al* (2011) | Myeloablative and reduced-intensity | Retrospective | Mismatched donor<br>Non-chronic phase disease<br>No splenectomy in men |

BuCy, busulfan and cyclophosphamide; HLA, human leucocyte antigen.
*These studies proposed predictive risk scores for myelofibrosis.

**Table II.** Risk models applied in the current study.

| Risk Model | Risk factors | Point per factor | Risk stratification |
|---|---|---|---|
| Lille | Hb <100 g/l | 1 | Low: 0 points |
| | WBC <4 or >30 × $10^9$/l | 1 | Intermediate: 1 points |
| | | | High: 2 points |
| Cervantes | Hb <100 g/l | 1 | Low: 0–1 points |
| | Circulating blasts≥1% | 1 | High: 2–3 points |
| | Constitutional symptoms | 1 | |
| IPSS | Age >65 years | 1 | Low: 0 points |
| | Hb <100 g/l | 1 | Intermediate-1: 1 points |
| | WBC >25 × $10^9$/l | 1 | Intermediate-2: 2 points |
| | Circulating blasts ≥ 1% | 1 | High: 3 points or more |
| | Constitutional symptoms | 1 | |
| Dynamic IPSS | Age >65 years | 1 | Low: 0 points |
| | Hb <100 g/l | 2 | Intermediate-1: 1 points |
| | WBC >25 × $10^9$/l | 1 | Intermediate-2: 3–4 points |
| | Circulating blasts ≥ 1% | 1 | High: 5–6 points |
| | Constitutional symptoms | 1 | |
| Age-adjusted dynamic IPSS | Hb <100 g/l | 2 | Low: 0 points |
| | WBC >25 × $10^9$/l | 1 | Intermediate-1: 1 points |
| | Circulating blasts ≥ 1% | 2 | Intermediate-2: 3–4 points |
| | Constitutional symptoms | 2 | High: >4 points |
| EBMT (adapted from Gratwohl *et al*, 2009) | Age of patient (years) | | Direct sum of points |
| | <20 | 0 | |
| | 20–40 | 1 | |
| | >40 | 2 | |
| | Disease stage* | | |
| | Early | 0 | |
| | Intermediate | 1 | |
| | Late | 2 | |
| | Time Interval from MF diagnosis to SCT† | | |
| | <18 months | 0 | |
| | >18 months | 1 | |
| | Donor type | | |
| | HLA-identical sibling | 0 | |
| | Unrelated donor | 1 | |
| | Donor-recipient sex combination | | |
| | All other | 0 | |
| | Female-male | 1 | |

IPSS, International Prognostic Scoring system; EBMT, European Blood and Marrow Transplantation Group; MF, myelofibrosis; SCT, stem cell transplantation; HLA, human leucocyte antigen.

*Disease stage was expressed as follows: early, when no peripheral blasts were detectable; intermediate, when 1–20% blasts were detectable; late, when peripheral blasts exceed 20%.

†According to the median of the current cohort.

Mann–Whitney *U*-test. To determine potential risk factors for survival in our cohort, a univariate Cox regression model was used to analyse the following factors: *JAK2* V617F-status, age, time from primary diagnosis to ASCT, time from myelofibrosis (or fibrotic transformation) to ASCT, transfusion dependency, massive splenomegaly, cytogenetics, white blood cell count (WBC) ≥25 × $10^9$/l, Hb ≤100 g/l, constitutional symptoms, peripheral blasts ≥1%, fibrosis grade, cytomegalovirus (CMV) serostatus of both patient and donor, sex of both patient and donor, donor and human leucocyte antigen (HLA)-match status, ABO-match status. Continuous variables were transformed into categorical ones either according to the median value of the current cohort or to other cut-off values when considered reasonable according to clinical expertise or related literature. In a second step, all pre-transplant factors with a *P*-value ≤0·05 were entered in a multivariate Cox regression model (Forward Wald) to identify the most significant independent prognostic factors affecting post-ASCT outcome.

**Table III.** Patient characteristics.

| Clinical characteristics | | *N* (%) or median (range) |
|---|---|---|
| Age (years) (*n* = 173) | | 57 (32–73) |
| | <57 years | 91 (53) |
| | ≥57 years | 82 (47) |
| Gender (*n* = 173) | Male | 86 (57) |
| | Female | 64 (43) |
| Diagnosis at SCT (*n* = 172) | PMF | 108 (63) |
| | Post PV/ET MF | 60 (35) |
| | Blast phase | 4 (2) |
| Time from myelofibrosis to SCT (months) (*n* = 170) | | 18 (1–272) |
| Time from MPN to SCT (months) (*n* = 166) | | 51 (1–363) |
| Blasts in peripheral blood, % (*n* = 144) | | 1 (0–35) |
| Transfusion dependency (*n* = 169) | No | 75 (44) |
| | Yes | 94 (56) |
| Massive splenomegaly (*n* = 93)* | No | 55 (59) |
| | Yes | 38 (41) |
| Splenectomy (*n* = 173) | No | 152 (89) |
| | Yes | 21 (11) |
| Cytogenetics (*n* = 115)* | Favourable | 9 (8) |
| | Normal | 63 (55) |
| | Others | 25 (22) |
| | Unfavourable | 18 (15) |
| Lille score (*n* = 173) | Low | 37 (21) |
| | Intermediate | 97 (56) |
| | High | 39 (23) |
| Cervantes score (*n* = 173) | Low | 43 (25) |
| | High | 130 (75) |
| IPSS score (*n* = 158) | Low | 3 (2) |
| | Intermediate-1 | 25 (16) |
| | Intermediate-2 | 50 (32) |
| | High | 80 (50) |
| Age-adapted IPSS score (*n* = 158)† | Low | 2 (1) |
| | Intermediate-1 | 18 (11) |
| | Intermediate-2 | 42 (27) |
| | High | 96 (61) |
| DIPSS score (*n* = 158) | Low | 3 (2) |
| | Intermediate-1 | 45 (29) |
| | Intermediate-2 | 91 (58) |
| | High | 19 (11) |
| aaDIPSS score (*n* = 135) | Low | 3 (2) |
| | Intermediate-1 | 25 (19) |
| | Intermediate-2 | 49 (36) |
| | High | 58 (43) |
| EBMT-Gratwohl risk model (*n* = 142) | Score 2 | 6 (4) |
| | Score 3 | 27 (19) |
| | Score 4 | 49 (35) |
| | Score 5 | 51 (36) |
| | Score 6 | 9 (6) |
| Fibrosis grade (*n* = 147)‡ | 0–1 | 15 (10) |
| | 2 | 35 (24) |
| | 3 | 97 (66) |
| *JAK2* V617F status (*n* = 150) | Pos | 102 (68) |
| | Neg | 48 (32) |
| CMV serostatus-Patient (*n* = 166) | Neg | 76 (46) |
| | Pos | 94 (54) |
| CMV serostatus-Donor (*n* = 166) | Neg | 72 (43) |
| | Pos | 94 (57) |
| Gender of donor (*n* = 173) | Male | 108 (62) |
| | Female | 65 (38) |
| ABO match status (*n* = 161) | Identical | 76 (47) |
| | Major MM | 32 (20) |
| | Major + minor MM | 9 (6) |
| | Minor MM | 44 (27) |
| Donor type (*n* = 173) | Sibling | 49 (28) |
| | Unrelated | 124 (72) |
| HLA-Match status (*n* = 173) | Matched | 139 (80) |
| | Mismatched | 34 (20) |

IPSS, International Prognostic Scoring system; DIPSS, dynamic IPSS; aa, age-adjusted; EBMT, European Blood and Marrow Transplantation Group; MF, myelofibrosis; SCT, stem cell transplantation; CMV, cytomegalovirus; HLA, human leucocyte antigen, MM, mismatch.

*Massive splenomegaly was defined as a spleen span of >15 cm under the left rib margin.

†Created by using age cut-off of 75 years instead of 65 years.

‡According to the European consensus on the grading of myelofibrosis (Thiele *et al*, 2005).

§Defined according to Hussein *et al* (2010): favourable: sole +9, 20q- or 13q-, unfavourable: complex karyotype or sole +8.

These factors were used to compose a scoring model after consideration of the individual weight of each factor reflected by hazard ratio (HR). Whenever reported, relapse risk and transplant-related mortality were expressed using cumulative incidence to account for competing risk. To evaluate the previously available scoring systems we estimated survival probabilities using the Kaplan–Meier method, where the log-rank test was used to compare survival curves. HRs and 95% Confidence intervals (95%CI) were reported using Cox regression method.

## Results

### *Defining an new scoring model*

The subgroup of 150 patients with a complete data set was evaluated. The median follow up was 21 months (range 0–130) and 5-year OS was 60% (95%CI 56–66%). Potential factors that could affect OS were included in an univariate Cox model where *JAK2* V617F wild-type (HR 1·83–95%CI 1·06–3·15; $P = 0·03$), age ≥57 years (HR 2·70–95%CI 1·52–4·78; $P = 0·001$), constitutional symptoms (HR 2·27–95%CI 1·14–4·53; $P = 0·02$), Fibrosis grade 3 (HR 2·85–95%CI 1·33–6·10; $P = 0·007$) and unrelated donor (HR 2·03–95%CI 1·02–4·04;

$P = 0{\cdot}05$) were statistically significant predictors (Table IV). As shown in Table IV, several other factors showed only trends affecting OS but did not reach statistical significance, like peripheral blasts, sex of donor, Donor-HLA-status. Due to the limited group and event numbers (48 events) these factors could not be included in the multivariate regression model. In the multivariate analysis only *JAK2* V617F wild-type, age ≥57 years and constitutional symptoms (HR 2·02; 2·43; 2·80

**Table IV.** Univariate analysis for OS.

| Potential factors | | HR | 95% CI | *P* value |
|---|---|---|---|---|
| *JAK2* V617F-status ($n = 150$) | Mutated | 1 | | |
| | Wild-type | 1·83 | 1·06–3·15 | 0·03 |
| Age ≥57 years ($n = 150$) | No | 1 | | |
| | Yes | 2·7 | 1·52–4·78 | 0·001 |
| Age ≥65 years ($n = 150$) | No | 1 | | |
| | Yes | 1·75 | 0·88–3·49 | 0·1 |
| Time from Myelofibrosis to SCT ≥18 months ($n = 148$) | No | 1 | | |
| | Yes | 0·9 | 0·51–1·50 | 0·6 |
| Time from Diagnosis to SCT ≥51 months ($n = 144$) | No | 1 | | |
| | Yes | 1·1 | 0·64–1·91 | 0·7 |
| Transfusion dependency ($n = 149$) | No | 1 | | |
| | Yes | 1·44 | 0·82–2·52 | 0·2 |
| Massive splenomegaly ($n = 92$) | No | 1 | | |
| | Yes | 0·78 | 0·39–1·58 | 0·4 |
| Splenectomy ($n = 150$) | No | 1 | | |
| | Yes | 1·67 | 0·60–4·62 | 0·3 |
| Cytogentics ($n = 99$) | Favourable | 1 | | 1·0 |
| | Normal | 0·96 | 0·29–3·26 | 1·0 |
| | Others | 0·95 | 0·25–3·58 | 0·9 |
| | Unfavourable | 1·12 | 0·28–4·51 | 0·9 |
| Leucocytes ≥25 × $10^9$/l ($n = 146$) | No | 1 | | |
| | Yes | 1·34 | 0·74–2·40 | 0·3 |
| Hb ≤100 g/l ($n = 147$) | No | 1 | | |
| | Yes | 1·65 | 0·86–3·14 | 0·1 |
| Constitutional symptoms ($n = 145$) | No | 1 | | |
| | Yes | 2·27 | 1·14–4·53 | 0·02 |
| Peripheral blasts ≥1% ($n = 146$) | No | 1 | | |
| | Yes | 1·87 | 0·93–3·74 | 0·08 |
| Fibrosis grade-3 ($n = 138$) | No | 1 | | |
| | Yes | 2·85 | 1·33–6·10 | 0·007 |
| CMV-serostatus of patient ($n = 150$) | neg | 1 | | |
| | pos | 1·51 | 0·85–2·68 | 0·2 |
| CMV-serostatus of donor ($n = 150$) | neg | 1 | | |
| | pos | 0·98 | 0·56–1·73 | 0·9 |
| Sex of patient ($n = 150$) | Male | 1 | | |
| | Female | 1·14 | 0·66–1·98 | 0·6 |
| Sex of donor ($n = 150$) | Male | 1 | | |
| | Female | 0·62 | 0·36–1·07 | 0·08 |
| Donor ($n = 150$) | Related | 1 | | |
| | Unrelated | 2·03 | 1·02–4·04 | 0·05 |
| HLA-match ($n = 150$) | Matched | 1 | | |
| | Mismatched | 1·53 | 0·84–2·78 | 0·2 |
| Donor-HLA-Status ($n = 150$) | Matched related | 1 | | 0·1 |
| | Matched unrelated | 1·89 | 0·92–3·9 | 0·09 |
| | Mismatched unrelated | 2·34 | 1·05–5·21 | 0·04 |
| ABO-match status ($n = 140$) | Identical | 1 | | 0·5 |
| | Major mismatch | 1·69 | 0·83–3·42 | 0·2 |
| | Major and minor mismatch | 1·45 | 0·43–4·90 | 0·6 |
| | Minor mismatch | 1·15 | 0·56–2·38 | 0·7 |

SCT, stem cell transplantation; CMV, cytomegalovirus; HLA, human leucocyte antigen.

respectively) were independent predictors of OS (Table V). In a next step, we stratified the cohort according to the presence of one, two or all of these factors and found the following HR values: 3·08; 4·70 and 16·61 ($P < 0·001$) respectively. Five-year OS was 90%, 74%, 51% for the presence of 0, 1 and 2 risk factors respectively, where the presence of 3 risk factors was associated with extremely poor life expectancy (1-year OS of 25%, median survival 5·5 months, range 3·0–8·1) (Table VI and Fig 1). Interestingly, this poor outcome was mainly caused by treatment-related mortality (data not shown). The four prognostic groups were well balanced for most of the examined clinical characteristics, engraftment data and graft-*versus*-host disease rate. The clinical variables that were unevenly distributed between the groups are outlined in Table VI.

**Table V.** Multivariate model.

| Potential factors | Hazard ratio | 95% Confidence interval | *P* value |
|---|---|---|---|
| *JAK2* V617F wild-type | 2·02 | 1·13–3·62 | 0·02 |
| Age ≥57 years | 2·43 | 1·33–4·44 | 0·004 |
| Constitutional symptoms | 2·8 | 1·30–6·02 | 0·008 |

### *Validation of previously available risk models*

Table VII and Fig 2 the Lille score had the best performance within all examined scoring models, where 5-year OS was 83%, 56% and 51% for low, intermediate and high-risk groups respectively (HR: 1; 3·18; 5·42 respectively). However it showed poor discrimination between intermediate and high-risk, especially in the long-term OS.

Patients in the high-risk category according to Cervantes score showed, as expected, a statistically non-significant trend toward reduced OS (57% vs. 70%, HR 1; 1·62 for high and low risk respectively). Regarding IPSS and DIPSS, only few of the

**Table VI.** Clinical characteristics and outcome of patients stratified by transplant risk score (only statistically significant differences are listed).

| | Low risk (no risk factors) $n = 19$ | Intermediate-1 risk (1 risk factor) $n = 53$ | Intermediate-2 risk (2 risk factors) $n = 57$ | High risk (3 risk factors) $n = 16$ | *P* value |
|---|---|---|---|---|---|
| Primary disease | | | | | |
| PMF | 8 (42) | 31 (60) | 38 (69) | 10 (67) | 0·05 |
| Post PV | 7 (37) | 10 (20) | 6 (11) | 0 (0) | |
| Post ET | 4 (21) | 10 (20) | 11 (20) | 5 (33) | |
| Tranfusion dependency | | | | | |
| No | 11 (58) | 28 (53) | 23 (40) | 3 (20) | 0·05 |
| Yes | 8 (42) | 25 (47) | 34 (60) | 13 (81) | |
| Lille score | | | | | |
| Low | 8 (42) | 11 (21) | 12 (21) | 0 (0) | 0·03 |
| Intermediate | 6 (32) | 32 (60) | 30 (53) | 11 (69) | |
| High | 5 (26) | 10 (19) | 15 (26) | 5 (31) | |
| Leucocytes ≥25 × $10^9$/l | | | | | |
| No | 10 (53) | 38 (73) | 45 (79) | 14 (88) | 0·09 |
| Yes | 9 (47) | 14 (27) | 12 (21) | 2 (13) | |
| Hb ≤ 100 g/l | | | | | |
| No | 9 (47) | 18 (35) | 18 (32) | 1 (6) | 0·04 |
| Yes | 10 (53) | 34 (65) | 39 (68) | 15 (94) | |
| Sex of patient | | | | | |
| Male | 10 (53) | 32 (60) | 27 (47) | 14 (88) | 0·04 |
| Female | 9 (47) | 21 (40) | 30 (53) | 2 (13) | |
| Grade of fibrosis | | | | | |
| 0–2 | 11 (58) | 20 (42) | 11 (21) | 1 (7) | 0·001 |
| 3 | 8 (42) | 28 (58) | 42 (79) | 14 (93) | |
| Overall survival | | | | | |
| HR (95% CI) | 1 | 3·1(0·7–13·6) | 4·7 (1·1–20·0) | 16·6 (3·7–74·5) | <0·001 |

HR, hazard ratio; 95%CI, 95% confidence interval.

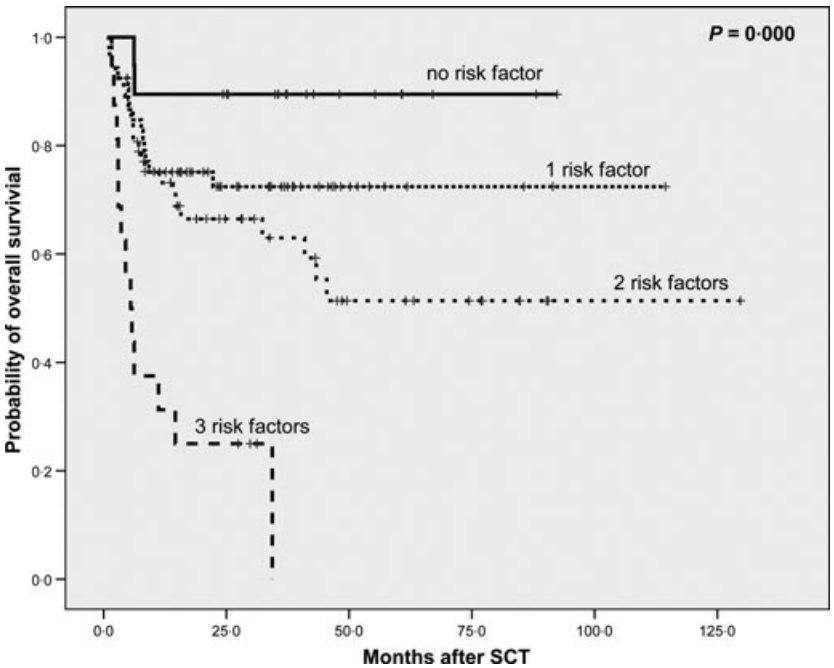


**Fig 1.** Survival based upon the presence of the following risk factors prior to autologous stem cell transplantation: *JAK2* wild-type, age ≥57 years and constitutional symptoms. SCT, stem cell transplantation.

**Table VII.** Performance of different scoring models in prediction of survival.

| Risk model | Stratification | 5-year OS (95% CI) | HR (95% CI) | *P* value |
|---|---|---|---|---|
| Lille | Low | 83 (71–95) | 1 | 0·003 |
| | Intermediate | 56 (46–66) | 3·18 (1·25–8·13) | 0·02 |
| | High | 51 (38–64) | 5·42 (2·02–14·56) | 0·001 |
| Cervantes | Low | 70 (57–83) | 1 | 0·2 |
| | High | 57 (50–64) | 1·62 (0·84–3·129) | |
| IPSS | Low+Intermediate-1 | 78 (65–91) | 1 | 0·001 |
| | Intermediate-2 | 78 (65–91) | 0·8 (0·3–2·37) | 0·7 |
| | High | 43 (32–54) | 2·7 (1·18–6·60) | 0·02 |
| IPSS (age-adjusted) | Low+Intermediate-1 | 85 (72–98) | 1 | 0·02 |
| | Intermediate-2 | 72 (59–85) | 1·73 (0·48–6·28) | 0·1 |
| | High | 49 (38–60) | 3·70 (1·15–11·97) | 0·000 |
| DIPSS | Low+Intermediate-1 | 70 (57–83) | 1 | 0·05 |
| | Intermediate-2 | 60 (50–70) | 1·52 (0·78–2·95) | 0·2 |
| | High | 47 (29–65) | 2·86 (1·23–6·63) | 0·02 |
| DIPSS (age-adapted) | Low+Intermediate-1 | 75 (62–88) | 1 | 0·000 |
| | Intermediate-2 | 78 (67–89) | 0·74 (0·29–1·90) | 0·5 |
| | High | 38 (26–50) | 2·65 (1·18–5·95) | 0·02 |
| EBMT risk model (Gratwohl *et al*, 2009) | Score 2 | 67 (37–97) | 1 | 0·8 |
| | Score 3 | 74 (61–87) | 0·81 (0·17–3·93) | |
| | Score 4 | 60 (48–72) | 1·32 (0·31–5·72) | |
| | Score 5 | 50 (30–70) | 1·17 (0·27–5·07) | |
| | Score 6 | 67 (42–92) | 1·20 (0·20–7·23) | |

IPSS, International Prognostic Scoring system; DIPSS, dynamic IPSS; EBMT, European Blood and Marrow Transplantation Group; HR, hazard ratio; 95% CI, 95% confidence interval.

studied patients were considered low risk prior to ASCT, so these small groups were combined with the corresponding intermediate-1 groups. IPSS, DIPSS and age-adjusted DIPSS (aaDIPSS) risk models could overall reliably distinguish patients with high risk but here the groups with lower risk status were not adequately discriminated (HR for low+inter-

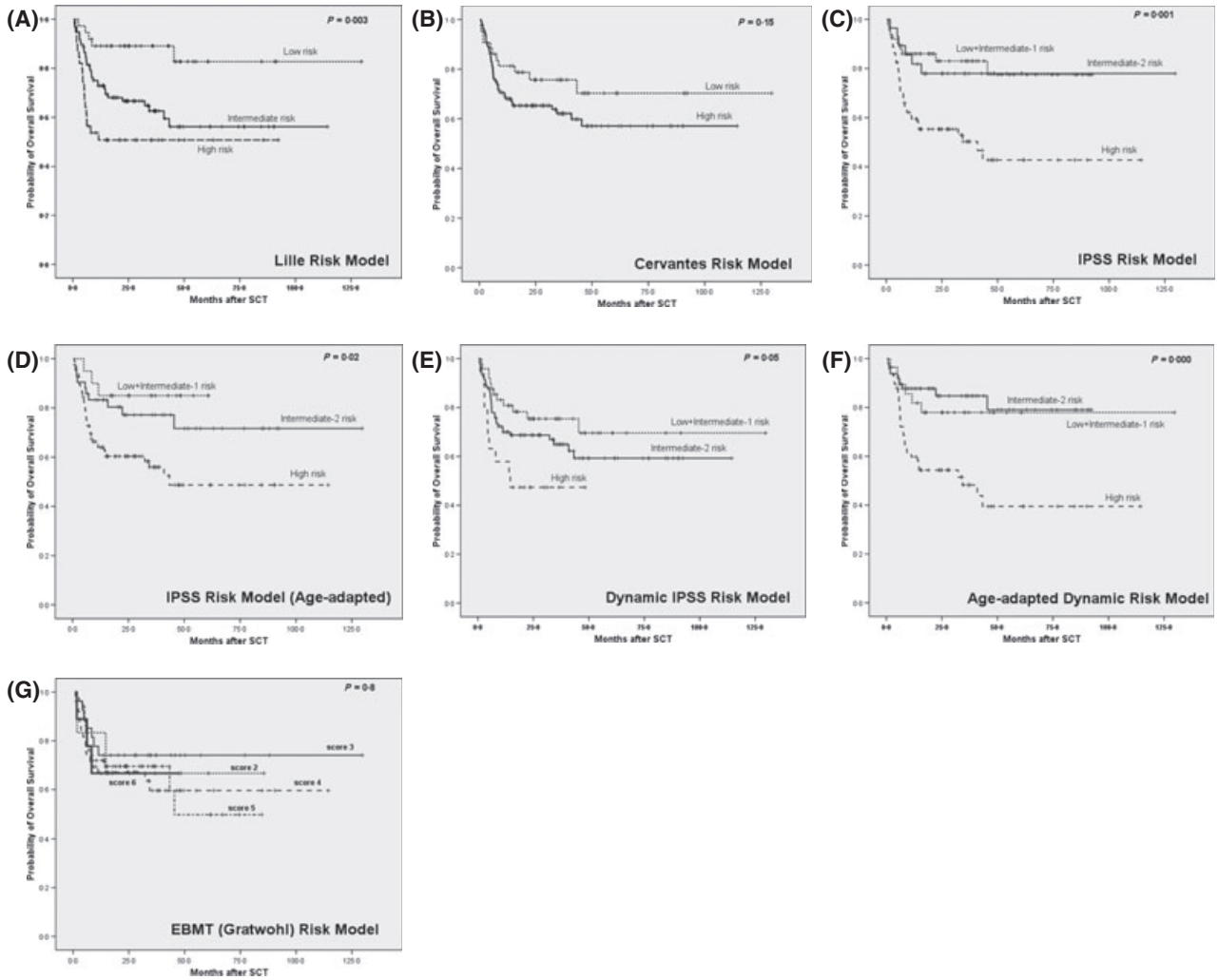


**Fig 2.** Performance of different scoring models in prediction of survival. IPSS International Prognostic Scoring system; EBMT, European Blood and Marrow Transplantation Group; SCT, stem cell transplantation.

mediate-1, intermediate-2 and high were 1·0 vs. 0·8 vs. 2·7 in IPSS; 1·0 vs. 1·52 vs. 2·86 in DIPSS and 1·0 vs. 0·74 vs. 2·65 in aaDIPSS respectively). Adjusting the age cut-off of the IPSS to 57 years according to the median of our cohort (Age-adapted IPSS) slightly improved the resolution (HR for low + intermediate-1, intermediate-2 and high were 1·0 vs. 1·73 vs. 3·70).

Finally, the modified EBMT score was unable to predict outcome after ASCT ($P = 0·8$).

## Discussion

Myelofibrosis is characterized by a highly variable life expectancy, which can range between few months to more than a decade (Dupriez *et al*, 1996). Risk stratification is needed to give a rough estimation of long-term survival, which can be implemented in patient consultations and choice of therapy. Patients are generally offered innovative and potentially curative therapies, such as ASCT, when they present with intermediate and high-risk disease and are transplant-eligible. However, the best tool to define patients with intermediate and high-risk is not yet certain.

The main aim of this study was to propose a scoring model for myelofibrosis patients who may undergo RIC-ASCT. To do so, we evaluated a homogenously treated cohort to identify independent predictive factors for OS. Here, advanced age, constitutional symptoms, as well as *JAK2*-status, remained in the final multivariate Cox regression model. The impact of *JAK2* V617F on survival after ASCT was recently reported by a study performed on a subgroup of this cohort (Kroger *et al*, 2009; Alchalby *et al*, 2010). To our knowledge, constitutional symptoms –defined as fever or night sweet persisting for more than 1 month or weight loss >10% of body weight in the last year- were not included in any statistical evaluation for OS after ASCT; nevertheless, its prognostic value for PMF in general is well established though not adequately understood (Dupriez *et al*, 1996; Cervantes *et al*, 1998, 2009). Constitutional symptoms are probably related to elevated cytokine levels, which have been shown recently to be associated with inferior OS and leukaemia-

free survival in PMF patients (Tefferi *et al*, 2011). It can be hypothesized that elevated cytokine levels reflected clinically by constitutional symptoms may intensify the damaging effect of the ASCT-associated cytokine storm. This interesting issue warrants further evaluation in future studies. Finally, the age of patients as an independent factor for OS in general and in the ASCT setting was shown previously (Kerbauy *et al*, 2007; Cervantes *et al*, 2009; Kroger *et al*, 2009). The resulting prognostic model divided the studied group into four risk categories with clearly distinct clinical behaviour post-ASCT and therefore may be used to identify different risk groups prior to ASCT.

Interestingly, some factors, such as grade of fibrosis and donor type, were statistically significant in the univariate analysis but their independence could not be confirmed in the multivariate model. On the other hand, cytogenetics, which were recently found to be prognostically important in myelofibrosis (Hussein *et al*, 2010) seem to play no role in the current cohort. A more controversial issue is splenectomy prior to ASCT, which has been shown in one study to be independently predictive for favourable survival in male patients (Robin *et al*, 2011) and in another to be significantly associated with relapse (Kroger *et al*, 2009). The current study did not identify any influence of splenectomy before ASCT on OS.

We are aware of some drawbacks in the current study, such as its retrospective nature, the small case volume for the tested hypothesis, the inclusion of cases with post-ET/PV myelofibrosis as well as the subjective element of constitutional symptoms, which may reduce the discrimination power of our score. However, taking into account that MPNs are rare diseases for which only a small proportion of patients are treated with ASCT, the reported group is one of the largest transplantation cohorts reported so far and, to our knowledge, the most homogenous one.

Recently, Bacigalupo *et al* (2010) reported a score before RIC-ASCT that included donor type, transfusion of more than 20 blood units and a spleen size >22 cm. In the present patient group, data regarding number of transfused blood units and degree of splenomegaly were not available to assess this interesting finding. However in the univariate analysis, no predictive value for OS could be attributed to transfusion need, irrespective of the number of units as well as the degree of splenomegaly. For the same reason, platelet count prior to ASCT could not be examined in the current analysis in spite of the potential prognostic relevance (Deeg *et al*, 2003; Kerbauy *et al*, 2007; Tam *et al*, 2009; Gangat *et al*, 2011).

In the second part of the study, the new proposed score was compared with the other most widely used scoring systems: Lille, Cervantes, IPSS, DIPSS and modified EBMT.

The Lille score is the scoring system most often applied in the literature of ASCT for myelofibrosis. Advanced Lille score was found to be an independent risk factor for OS in a report on standard-dose conditioning ASCT (Deeg *et al*, 2003) and for disease-free survival in another with RIC-ASCT (Kroger *et al*, 2009). However both reports showed that, although this model can reliably delineate the low-risk group, it fails to discriminate between intermediate and high-risk categories. This shortcoming is not confined to ASCT cohorts, as it has been reported in groups of patients treated without ASCT in different extents (Tefferi *et al*, 2007; Morel & Duhamel, 2010). This was also the case in the current study where the difference in terms of OS between the two groups was rather small: 51% vs. 56%. An advanced Cervantes score similarly showed the expected trend toward an inferior OS, which was not statistically significant. Remarkably, Cervantes score was originally proposed for patients younger than 55 years, which is actually a limitation for its application in RIC-ASCT patients. Both Lille and Cevantes scores are age-independent, which may contribute to their weakness in the transplant setting. In contrast, the IPSS model combines factors from both previously mentioned models but adds the age with a cut-off of 65 years. It could be expected that this model would provide a better discrimination between the risk groups prior to ASCT. This was again not the case in the current analysis, where the three lower risk groups were not distinguishable in contrast to the high-risk group, which was relatively well delineated. Due to our previous knowledge (Kroger *et al*, 2009), that the effect of age is distributed differently in a transplantation cohort, where a cut-off point of 57 years (according to the current cohort) may reflect better the survival difference, we applied IPSS again using the new age-limit but could not achieve a better discrimination. Notably, the Lille, Cervantes and IPSS scoring systems were assigned to be applied at diagnosis, but not at other later time points, which is a limitation for their use prior to ASCT. DIPSS was introduced recently to apply at every time point during disease evolution. As such, it may be suitable to use this system prior to ASCT. Interestingly, this score involves the same risk factors used for IPSS but applies more weight for anaemia, which was not predictive for survival in our cohort and in turn explains why the outcome did not improve compared to IPSS. Analysing a subgroup of young patients (age limit 65 years) after classification according to the aaDIPSS, we detected a sharper delineation of the high-risk group but the groups with lower risk were still not distinguishable.

The transplant score initially proposed by Gratwohl *et al* (1998) for patients with chronic myeloid leukaemia and a decade thereafter for haematological disease in general (Gratwohl *et al*, 2009) was shown to be predictive for transplant outcome. Interestingly, in the work published in 2009 in which the applicability of this score in diverse haematological malignancies was evaluated, patients with myelofibrosis were not included. Questions, such as the best measure for disease stage and the most suitable cut-off for time interval between diagnosis and ASCT, are in our opinion unclear due to the heterogenous and largely unclear nature of this disease. However, application of this score according to our interpretation in this cohort did not offer any predictive value.

As a result, we believe that our new risk model may be used along with the Lille score for patient consultation prior to RIC-ASCT and also to identify patients with lower risk groups characterized with high cure potential from those high-risk

patients, where the overall mortality may warrant therapy adjustment or where enrollment in prospective studies with investigational protocols is more appropriate. Furthermore, it would probably be of interest to explore whether the administration of JAK2-inhibitors prior to ASCT, which were found to alleviate constitutional symptoms and, to a lesser degree, reduce *JAK2* allele burden, may consequently improve transplant outcome. Undoubtedly, this new model needs further validation and confirmation through independent RIC-ASCT cohorts.

## Acknowledgements

The authors would like to thank Prof. Jürgen Berger from the Department of Medical Biometry and Epidemiology, University Hospital Hamburg-Eppendorf for the statistical consultation and advice.

HA collected and analysed patients' data, performed the statistical analysis, interpreted results and wrote the manuscript; DR.Y collected patients' data; TZ performed statistical analysis; GK, EH, MB, RS, WB, HMK and GB submitted patients' data; FA, UB and ARZ analysed data and interpreted results; NK designed the study, and wrote the manuscript; and all authors reviewed and approved the manuscript.

## Conflict of interest statements

Authors have no relevant disclosures.

## References

Alchalby, H., Badbaran, A., Zabelina, T., Kobbe, G., Hahn, J., Wolff, D., Bornhauser, M., Thiede, C., Baurmann, H., Bethge, W., Hildebrandt, Y., Bacher, U., Fehse, B., Zander, A.R. & Kroger, N. (2010) Impact of JAK2V617F mutation status, allele burden, and clearance after allogeneic stem cell transplantation for myelofibrosis. *Blood*, **116,** 3572–3581.

Bacigalupo, A., Soraru, M., Dominietto, A., Pozzi, S., Geroldi, S., Van Lint, M.T., Ibatici, A., Raiola, A.M., Frassoni, F., De Stefano, F., Verdiani, S., Casarino, L. & Barosi, G. (2010) Allogeneic hemopoietic SCT for patients with primary myelofibrosis: a predictive transplant score based on transfusion requirement, spleen size and donor type. *Bone Marrow Transplantation*, **45,** 458–463.

Ballen, K.K., Shrestha, S., Sobocinski, K.A., Zhang, M.J., Bashey, A., Bolwell, B.J., Cervantes, F., Devine, S.M., Gale, R.P., Gupta, V., Hahn, T.E., Hogan, W.J., Kroger, N., Litzow, M.R., Marks, D.I., Maziarz, R.T., McCarthy, P.L., Schiller, G., Schouten, H.C., Roy, V., Wiernik, P.H., Horowitz, M.M., Giralt, S.A. & Arora, M. (2010) Outcome of transplantation for myelofibrosis. *Biol Blood Marrow Transplant*, **16,** 358–367.

Cervantes, F., Barosi, G., Demory, J.L., Reilly, J., Guarnone, R., Dupriez, B., Pereira, A. & Montserrat, E. (1998) Myelofibrosis with myeloid metaplasia in young individuals: disease characteristics, prognostic factors and identification of risk groups. *British Journal of Haematology*, **102,** 684–690.

Cervantes, F., Passamonti, F. & Barosi, G. (2008) Life expectancy and prognostic factors in the classic BCR/ABL-negative myeloproliferative disorders. *Leukemia*, **22,** 905–914.

Cervantes, F., Dupriez, B., Pereira, A., Passamonti, F., Reilly, J.T., Morra, E., Vannucchi, A.M., Mesa, R.A., Demory, J.L., Barosi, G., Rumi, E. & Tefferi, A. (2009) New prognostic scoring system for primary myelofibrosis based on a study of the International Working Group for Myelofibrosis Research and Treatment. *Blood*, **113,** 2895–2901.

Deeg, H.J., Gooley, T.A., Flowers, M.E., Sale, G.E., Slattery, J.T., Anasetti, C., Chauncey, T.R., Doney, K., Georges, G.E., Kiem, H.P., Martin, P.J., Petersdorf, E.W., Radich, J., Sanders, J.E., Sandmaier, B.M., Warren, E.H., Witherspoon, R.P., Storb, R. & Appelbaum, F.R. (2003) Allogeneic hematopoietic stem cell transplantation for myelofibrosis. *Blood*, **102,** 3912–3918.

Dupriez, B., Morel, P., Demory, J.L., Lai, J.L., Simon, M., Plantier, I. & Bauters, F. (1996) Prognostic factors in agnogenic myeloid metaplasia: a report on 195 cases with a new scoring system. *Blood*, **88,** 1013–1018.

Gangat, N., Caramazza, D., Vaidya, R., George, G., Begna, K., Schwager, S., Van Dyke, D., Hanson, C., Wu, W., Pardanani, A., Cervantes, F., Passamonti, F. & Tefferi, A. (2011) DIPSS Plus: a refined Dynamic International Prognostic Scoring System for Primary Myelofibrosis That Incorporates Prognostic Information From Karyotype, Platelet Count, and Transfusion Status. *Journal of Clinical Oncology*, **29,** 392–397.

Gratwohl, A., Hermans, J., Goldman, J.M., Arcese, W., Carreras, E., Devergie, A., Frassoni, F., Gahrton, G., Kolb, H.J., Niederwieser, D., Ruutu, T., Vernant, J.P., de Witte, T. & Apperley, J. (1998) Risk assessment for patients with chronic myeloid leukaemia before allogeneic blood or marrow transplantation. Chronic leukemia working party of the European group for blood and marrow transplantation. *Lancet*, **352,** 1087–1092.

Gratwohl, A., Stern, M., Brand, R., Apperley, J., Baldomero, H., de Witte, T., Dini, G., Rocha, V., Passweg, J., Sureda, A., Tichelli, A. & Niederwieser, D. (2009) Risk score for outcome after allogeneic hematopoietic stem cell transplantation: a retrospective analysis. *Cancer*, **115,** 4715–47260.

Guardiola, P., Anderson, J.E., Bandini, G., Cervantes, F., Runde, V., Arcese, W., Bacigalupo, A., Przepiorka, D., O'Donnell, M.R., Polchi, P., Buzyn, A., Sutton, L., Cazals-Hatem, D., Sale, G., de Witte, T., Deeg, H.J. & Gluckman, E. (1999) Allogeneic stem cell transplantation for agnogenic myeloid metaplasia: a European Group for Blood and Marrow Transplantation, Societe Francaise de Greffe de Moelle, Gruppo Italiano per il Trapianto del Midollo Osseo, and Fred Hutchinson Cancer Research Center Collaborative Study. *Blood*, **93,** 2831–2838.

Hussein, K., Pardanani, A.D., Van Dyke, D.L., Hanson, C.A. & Tefferi, A. (2010) International Prognostic Scoring System-independent cytogenetic risk categorization in primary myelofibrosis. *Blood*, **115,** 496–499.

Kerbauy, D.M., Gooley, T.A., Sale, G.E., Flowers, M.E., Doney, K.C., Georges, G.E., Greene, J.E., Linenberger, M., Petersdorf, E., Sandmaier, B.M., Scott, B.L., Sorror, M., Stirewalt, D.L., Stewart, F.M., Witherspoon, R.P., Storb, R., Appelbaum, F.R. & Deeg, H.J. (2007) Hematopoietic cell transplantation as curative therapy for idiopathic myelofibrosis, advanced polycythemia vera, and essential thrombocythemia. *Biol Blood Marrow Transplant*, **13,** 355–365.

Kroger, N., Zabelina, T., Schieder, H., Panse, J., Ayuk, F., Stute, N., Fehse, N., Waschke, O., Fehse, B., Kvasnicka, H.M., Thiele, J. & Zander, A. (2005) Pilot study of reduced-intensity conditioning followed by allogeneic stem cell transplantation from related and unrelated donors in patients with myelofibrosis. *British Journal Haematology*, **128,** 690–697.

Kroger, N., Holler, E., Kobbe, G., Bornhauser, M., Schwerdtfeger, R., Baurmann, H., Nagler, A., Bethge, W., Stelljes, M., Uharek, L., Wandt, H., Burchert, A., Corradini, P., Schubert, J., Kaufmann, M., Dreger, P., Wulf, G.G., Einsele, H., Zabelina, T., Kvasnicka, H.M., Thiele, J., Brand, R., Zander, A.R., Niederwieser, D. & de Witte, T.M. (2009) Allogeneic stem cell transplantation after reduced-intensity conditioning in patients with myelofibrosis: a prospective, multicenter study of the Chronic Leukemia Working Party of the European Group for Blood and Marrow Transplantation. *Blood*, **114,** 5264–5270.

Morel, P. & Duhamel, A. (2010) Comparison of prognostic scoring systems in primary myelofibrosis. *Blood*, **115,** 745.

Passamonti, F., Cervantes, F., Vannucchi, A.M., Morra, E., Rumi, E., Cazzola, M. & Tefferi, A. (2010) Dynamic International Prognostic Scoring

System (DIPSS) predicts progression to acute myeloid leukemia in primary myelofibrosis. *Blood*, **116,** 2857–2858.

Robin, M., Tabrizi, R., Mohty, M., Furst, S., Michallet, M., Bay, J.O., Cahn, J.Y., De Coninck, E., Dhedin, N., Bernard, M., Rio, B., Buzyn, A., Huynh, A., Bilger, K., Bordigoni, P., Contentin, N., Porcher, R., Socie, G. & Milpied, N. (2011) Allogeneic haematopoietic stem cell transplantation for myelofibrosis: a report of the Societe Francaise de Greffe de Moelle et de Therapie Cellulaire (SFGM-TC). *British Journal Haematology*, **152,** 331–339.

Tam, C.S., Kantarjian, H., Cortes, J., Lynn, A., Pierce, S., Zhou, L., Keating, M.J., Thomas, D.A. & Verstovsek, S. (2009) Dynamic model for predicting death within 12 months in patients with primary or post-polycythemia vera/essential thrombocythemia myelofibrosis. *Journal of Clinical Oncology*, **27,** 5587–5593.

Tefferi, A., Huang, J., Schwager, S., Li, C.Y., Wu, W., Pardanani, A. & Mesa, R.A. (2007) Validation and comparison of contemporary prognostic models in primary myelofibrosis: analysis based on 334 patients from a single institution. *Cancer*, **109,** 2083–2088.

Tefferi, A., Vaidya, R., Caramazza, D., Finke, C., Lasho, T. & Pardanani, A. (2011) Circulating Interleukin (IL)-8, IL-2R, IL-12, and IL-15 levels are independently prognostic in primary myelofibrosis: A comprehensive cytokine profiling Study. *Journal of Clinical Oncology*, **29,** 1356–1363.

Thiele, J., Kvasnicka, H.M., Facchetti, F., Franco, V., van der Walt, J. & Orazi, A. (2005) European consensus on grading bone marrow fibrosis and assessment of cellularity. *Haematologica*, **90,** 1128–1132.