

MYELOID NEOPLASIA

Integrating clinical features and genetic lesions in the risk assessment of patients with chronic myelomonocytic leukemia

Chiara Elena,^{1,2} Anna Galli,² Esperanza Such,³ Manja Meggendorfer,⁴ Ulrich Germing,⁵ Ettore Rizzo,⁶ Jose Cervera,³ Elisabetta Molteni,¹ Annette Fasan,⁴ Esther Schuler,⁵ Ilaria Ambaglio,² Maria Lopez-Pavia,³ Silvia Zibellini,² Andrea Kuendgen,⁵ Erica Travaglio,² Reyes Sancho-Tello,⁷ Silvia Catricalà,² Ana I. Vicente,⁸ Torsten Haferlach,⁴ Claudia Haferlach,⁴ Guillermo F. Sanz,³ Luca Malcovati,^{1,2,*} and Mario Cazzola^{1,2,*}

¹Department of Molecular Medicine, University of Pavia, Pavia, Italy; ²Department of Hematology Oncology, Istituto di Ricovero e Cura a Carattere Scientifico Fondazione Policlinico San Matteo, Pavia, Italy; ³Department of Hematology, Hospital Universitario y Politecnico La Fe, Valencia, Spain; ⁴Munich Leukemia Laboratory GmbH, Munich, Germany; ⁵Department of Hematology, Oncology and Clinical Immunology, University Hospital Düsseldorf, Düsseldorf, Germany; ⁶Department of Electrical, Computer and Biomedical Engineering, University of Pavia, Pavia, Italy; ⁷Department of Hematology, Hospital Arnau de Vilanova, Valencia, Spain; and ⁸Department of Hematology, Hospital de La Ribera, Alzira, Spain

Key Points

- Risk assessment is crucial in patients with CMML because survival may range from a few months to several years.
- Integrating clinical features, morphology, and genetic lesions significantly improves risk stratification in CMML.

Chronic myelomonocytic leukemia (CMML) is a myelodysplastic/myeloproliferative neoplasm with variable clinical course. To predict the clinical outcome, we previously developed a CMML-specific prognostic scoring system (CPSS) based on clinical parameters and cytogenetics. In this work, we tested the hypothesis that accounting for gene mutations would further improve risk stratification of CMML patients. We therefore sequenced 38 genes to explore the role of somatic mutations in disease phenotype and clinical outcome. Overall, 199 of 214 (93%) CMML patients carried at least 1 somatic mutation. Stepwise linear regression models showed that these mutations accounted for 15% to 24% of variability of clinical phenotype. Based on multivariable Cox regression analyses, cytogenetic abnormalities and mutations in *RUNX1*, *NRAS*, *SETBP1*, and *ASXL1* were independently associated with overall survival (OS). Using these parameters, we defined a genetic score that identified 4 categories with significantly different OS and cumulative

incidence of leukemic evolution. In multivariable analyses, genetic score, red blood cell transfusion dependency, white blood cell count, and marrow blasts retained independent prognostic value. These parameters were included into a clinical/molecular CPSS (CPSS-Mol) model that identified 4 risk groups with markedly different median OS (from >144 to 18 months, hazard ratio [HR] = 2.69) and cumulative incidence of leukemic evolution (from 0% to 48% at 4 years, HR = 3.84) ($P < .001$). The CPSS-Mol fully retained its ability to risk stratify in an independent validation cohort of 260 CMML patients. In conclusion, integrating conventional parameters and gene mutations significantly improves risk stratification of CMML patients, providing a robust basis for clinical decision-making and a reliable tool for clinical trials. (*Blood*. 2016;128(10):1408-1417)

Introduction

Chronic myelomonocytic leukemia (CMML) is a myeloid neoplasm that shares both dysplastic and proliferative features.^{1,2} The French-American-British (FAB) classification distinguished 2 subtypes of CMML, a myelodysplastic and a myeloproliferative variant, based on a leukocyte count less than or equal to $13 \times 10^9/L$, respectively.³ The World Health Organization (WHO) in its classification of myeloid neoplasms included CMML in the category of myelodysplastic/myeloproliferative neoplasms (MDS/MPN),⁴ and in the 2016 revision recognized the dysplastic and proliferative types, and differentiated 3 groups of CMML according to the percentage of blasts in bone marrow (BM).⁵

The hematologic and morphologic features of CMML are highly heterogeneous, varying from predominantly myelodysplastic to mainly proliferative features, and the clinical course is extremely variable, with

wide differences in survival and risk of evolution into acute myeloid leukemia (AML).⁶ Different prognostic scoring systems were developed in the attempt to stratify individual patient risk.⁷⁻¹² We recently developed and validated a CMML-specific Prognostic Scoring System (CPSS),¹³ which combined cytogenetic abnormalities,¹⁴ disease subtype according to FAB and WHO classifications, and red blood cell (RBC) transfusion dependency, and stratified patients into 4 different risk groups with significantly different survival and risk of AML evolution.

Recurrent somatic mutations have been identified in a high proportion of patients with CMML. Mutant genes encode signaling molecules (*NRAS*, *KRAS*, *CBL*, and *JAK2*), epigenetic regulators (*TET2*, *IDH1*, *IDH2*, *DNMT3A*, *ASXL1*, and *EZH2*), splicing factors (*SRSF2*, *SF3B1*, *ZRSR2*, and *U2AF1*), and transcription factors (*RUNX1*).¹⁵⁻²⁰ More

Submitted May 2, 2016; accepted June 27, 2016. Prepublished online as *Blood* First Edition paper, July 6, 2016; DOI 10.1182/blood-2016-05-714030.

*L.M. and M.C. contributed equally to this study.

The online version of this article contains a data supplement.

There is an Inside *Blood* Commentary on this article in this issue.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

© 2016 by The American Society of Hematology

recently, mutations in *SETBP1*, *ETNK1*, and *CSF3R* have been identified in patients with MDS/MPN, including CMML.^{18,19,21} *TET2* and *SRSF2* mutations are the most prevalent in CMML,^{16,22} and the co-occurrence of mutations in these 2 genes has been found to be highly specific for myeloid neoplasms with monocytosis.²³ Associations between somatic mutations and disease phenotype have been suggested, with mutations in signaling pathways and the DNA methylation pathway being predominant in myeloproliferative and myelodysplastic CMML subtypes, respectively.²⁴

In addition, selected mutated genes have been recognized to provide useful prognostic information.^{20,22,25,26} In particular, mutations in *ASXL1* were consistently reported to have an independent prognostic value,^{16,26,27} and have been recently demonstrated to be of additive value to a clinically-based risk assessment.^{20,26} Conversely, conflicting results have been provided by the studies published so far on the prognostic relevance of other gene mutations.^{22,26,28}

In this study, we performed a comprehensive mutation analysis of genes implicated in myeloid malignancies in 2 independent cohorts of well-annotated patients with CMML from different institutions, with the aim to dissect the relationship between genotype and disease phenotype and to integrate mutation analysis into a clinical/molecular prognostic scoring system that may improve individual patient risk assessment.

Patients and methods

Patients' characteristics and clinical procedures

This study included 2 cohorts of CMML patients diagnosed according to the criteria of the 2008 WHO classification of myeloid neoplasms⁴ and its 2016 revision⁵: a learning cohort, in which we conducted investigations aimed at defining the variables to be included in the prognostic model, and a validation cohort, in which we evaluated whether the prognostic value of the scoring system was confirmed. The learning cohort consisted of 214 CMML patients diagnosed at different institutions: Istituto di Ricovero e Cura a Carattere Scientifico (IRCCS) Fondazione Policlinico San Matteo, Pavia, Italy; Hospital Universitario y Politécnico La Fe, Valencia, Spain; University Hospital Düsseldorf, Düsseldorf, Germany; Hospital Arnau de Vilanova, Valencia, Spain; and the Hospital de La Ribera, Alzira, Spain. The validation cohort consisted of 260 CMML patients diagnosed at the MLL Munich Leukemia Laboratory, Munich, Germany, and at the Fondazione IRCCS Policlinico San Matteo, Pavia, Italy.

This study was approved by the Ethics Committee of the Fondazione IRCCS Policlinico San Matteo and other local Institutional Review Boards. The procedures followed were in accordance with the Declaration of Helsinki of 1975, as revised in 2000, and samples were obtained after patients provided written informed consent.

Clinical and hematologic features of patients included in the learning and validation cohorts at the time of biological sampling are reported in Table 1. CMML-MD and CMML-MP were defined according to FAB criteria,³ whereas subclassification into CMML-0, CMML-1, and CMML-2, and evolution into AML were defined according to the 2016 WHO criteria.⁵ Quantitative enumeration of myeloblasts, monocytes, and their precursors was performed using established consensus criteria.^{29,30}

RBC-transfusion dependency and sex-specific hemoglobin thresholds were defined as previously reported.^{13,31,32} Cytogenetic risk category was defined according to the Spanish cytogenetic risk stratification,^{13,14} and CPSS risk was estimated according to Such et al (see supplemental Table 1, available on the Blood Web site).¹³

Sample collection and cell separation

Mononuclear cells were separated from BM samples by standard density gradient centrifugation, and granulocytes were isolated from peripheral blood, as previously described.³³ Peripheral T lymphocytes were purified by

immunomagnetic separation (Miltenyi Biotec, Bergisch Gladbach, Germany) according to the manufacturer's recommendations in 65 patients, and used as control tissue. Genomic DNA was obtained from tumor and control cells by following standard protocols for human tissue.

Mutation analysis

Mutation analysis was performed on DNA from peripheral blood granulocytes or BM mononuclear cells. A TruSeq Custom Amplicon panel (TSCA; Illumina, San Diego, CA) targeting complete coding exons and their adjacent splice junctions from 38 genes was designed using Illumina Design Studio software. The genes included in the panel were selected based on the available evidence in myeloid neoplasms (supplemental Table 2). The TSCA panel consisted of 886 amplicons, 425 bp in length, for a total of 205 kb targeted DNA. Dual-barcoded TSCA libraries were created from 250 ng of high quality DNA according to the manufacturer's protocol. Libraries were multiplexed and underwent 2 × 250-bp paired-end sequencing on a MiSeq sequencing system using the MiSeq Reagent Kit version 3 (Illumina). Mutational analysis of low performer regions (ie, regions with inadequate coverage, <100 reads) was carried out using Nextera XT Samples Preparation Kit (Illumina), and sequencing reactions were performed using the MiSeq version 2 (2 × 150 bp) chemistry. The resulting average depth of coverage for the 886 amplicons was 739.

Sequence reads were initially aligned to the human genome (GRCh37/hg19) using the Burrows-Wheeler Aligner.³⁴ The Genome Analysis Toolkit was then used to cleanup reads and make alignment data more reliable for the variant calling (Genome Analysis Toolkit data cleanup best practice).³⁵ Single nucleotide variants and small insertions and deletions were identified by UnifiedGenotyper.³⁵ Functionally annotated variants were then filtered based on the information retrieved from public databases (dbSNP, 1000genome, and ESP6500), the expected germ line allele frequency, as previously reported.¹⁷ The information obtained from patient-specific control tissue was used to confirm filtering strategy when available. The remaining variants were considered as candidate somatic mutations, and were finally tagged as oncogenic, based on the information derived from the literature and on in silico prediction effect, as previously described.¹⁷ Variant allele frequency was calculated as the number of variant reads divided by the total reads.

Statistical analysis

Survival analyses were performed with the Kaplan-Meier method. Multivariable survival analyses were performed by means of Cox proportional hazards regression. The cumulative incidence of progression to AML was estimated with a competing risk approach, considering death for any cause as a competing event, and the effect of quantitative covariates was estimated by Fine-Gray regression model.³⁶ All analyses accounted for left censoring of the observations at the time of mutation assessment. Any observation was censored before starting any disease-modifying therapy. The comparison of models with different types of covariates was carried out using Akaike's information criterion and Harrell's C concordance index.

The R^2 coefficient was calculated in order to evaluate the goodness of fit of linear regression models. This coefficient is the ratio between the explained variation and the total variation of the model: the better the linear regression fits the data, the closer the value of R^2 is to 1. Statistical analyses were performed using Stata 12.1 (StataCorp LP) software.

Results

Mutation spectrum and correlations between genotype and disease phenotype in CMML

Ninety-three percent of patients (199/214) showed at least 1 somatic mutation (median number per patient = 2, range, 0-7). The most frequently mutated genes were *TET2* (44.4% of patients), *SRSF2* (38.8%), *ASXL1* (37.0%), *NRAS* (11.7%), *KRAS* (8.9%), *SETBP1* (8.9%), *CBL* (8.4%), *RUNX1* (7.9%), *JAK2* (7.0%), *EZH2* (7.0%), *SF3B1* (5.6%), *IDH2* (5.6%), *U2AF1* (4.2%), and *ZRSR2* (4.2%)

Table 1. Clinical and hematologic features of patients with CMML included in the study

Variable	Learning cohort	Validation cohort
Patients, n	214	260
Follow-up in months, median (range)	22 (0.4-212.6)	17 (0.2-155)
Age in years, median (range)	72 (28-99)	74 (22-93)
Males, n (%)	151 (71)	190 (73)
FAB subtype, n (%)		
CMML-MD	129 (60)	148 (57)
CMML-MP	85 (40)	112 (43)
WHO subtype, n (%)		
CMML-0	138 (64)	112 (44)
CMML-1	38 (18)	74 (28)
CMML-2	38 (18)	74 (28)
WBC count, $\times 10^9/L$, median (range)	9.8 (1.2-126)	11.0 (2.0-87.0)
Absolute monocyte count, $\times 10^9/L$, median (range)	2.06 (0.97-34.01)	3.37 (1.02-39.95)
Hemoglobin level (g/dL), median (range)	11.6 (6-16.6)	11.5 (5.9-17)
Platelet count ($\times 10^9/L$), median (range)	124 (4-943)	103 (3-1385)
BM blasts* (%), median (range)	3 (1-18)	6 (0-19)
Chromosomal abnormalities, n (%)	44 (23)	58 (22)
Median survival (mo)	48.4	51.2
CPSS cytogenetic risk group,† n (%)		
Low	150 (79)	218 (84)
Intermediate	21(11)	13 (5)
High	19 (10)	29 (11)
CPSS risk group,‡ n (%)		
Low	85 (46)	84 (32)
Intermediate-1	49 (27)	102 (39)
Intermediate-2	46 (25)	66 (25)
High	4 (2)	8 (3)

CMML-MD, myelodysplastic CMML; CMML-MP, myeloproliferative CMML; WBC, white blood cell.

*BM blasts were analyzed as a percentage of BM nuclear cells.

†Cytogenetic abnormalities were categorized according to Such et al.^{13,14}

‡CPSS risk was estimated according to Such et al.¹³

(Figure 1). A significant association was found between mutations in *TET2* and RNA splicing factors ($P = .039$), 41 of 214 CMML patients (19%) showing co-occurrence of *TET2* and *SRSF2* mutations.

We analyzed the distribution of somatic mutations according to dysplastic or proliferative subtypes and WHO blast categories, and found that mutations in *JAK2*, *NRAS*, and *SETBP1* were significantly associated with CMML-MP (P values from .03 to $< .001$), whereas *TET2* and *SF3B1* mutations were associated with CMML-MD ($P = .007$ and $P = .03$, respectively) (supplemental Table 3; supplemental Figure 1A-B). Conversely, no significant association was found between mutation pattern and WHO category.

In order to estimate what fraction of the variability of clinical phenotype can be explained by genotype, we applied stepwise linear regression models using the clinical variables of interest (BM blasts, WBC count, and hemoglobin level) as the dependent variables and genetic lesions (point mutations and cytogenetic alterations) as predictors. To this aim, the R^2 coefficient, that is, the ratio between the explained variation and the total variation of the model, was calculated. BM blasts were independently predicted by the high CPSS cytogenetic risk group ($P = .002$) and *RUNX1* mutations ($P = .004$). Mutations in *SETBP1* ($P < .001$), *NRAS* ($P = .01$), and *RUNX1* ($P = .011$) were independently associated with WBC count, whereas *SF3B1* mutations were the only independent genetic predictor of hemoglobin level ($P = .032$). Genetic lesions accounted for 0.15, 0.20, and 0.24 of the variability of hemoglobin, WBC count, and BM blasts, respectively (Figure 2A-C). Comparable results were obtained for categorical variables, ie, FAB category (Pseudo R^2 0.20) and WHO subtype (Pseudo R^2 0.17) (supplemental Figure 2A-B).

Prognostic value of genetic lesions in CMML

The number of mutations per patient inversely correlated with overall survival (OS) (hazard ratio [HR] = 1.23, $P = .001$) (supplemental Figure 3). In univariable analysis, *ASXL1* nonsense/frameshift mutations (HR = 1.74, $P = .013$), and mutations in *EZH2* (HR = 2.07, $P = .023$), *NRAS* (HR = 2.44, $P = .001$), *RUNX1* (HR = 3.24, $P < .001$), and *SETBP1* (HR = 2.49, $P = .002$) significantly affected OS (supplemental Table 4). No significant effect of variant allele frequency on the prognostic value of mutated genes was noticed (P values ranging from .75 to .16).

In order to investigate the additive value of somatic mutations, we first performed multivariable Cox regression including genetic variables. The variables that retained independent prognostic value on OS were CPSS cytogenetic risk groups (HR = 1.54, $P = .001$), and mutations in *RUNX1* (HR = 2.32, $P = .016$), *NRAS* (HR = 2.19, $P = .009$), *SETBP1* (HR = 2.00, $P = .04$), and *ASXL1* (HR = 1.77, $P = .022$) (supplemental Figure 4A-D). Based on regression coefficients, we defined a CMML-specific genetic score that was able to identify 4 groups with significantly different OS and cumulative incidence of AML evolution (HR = 1.87 and HR = 2.62 respectively, $P < .001$) (Table 2; supplemental Figure 5A-B). The Akaike information criterion (AIC) showed that this genetic score performed better than the original CPSS cytogenetic risk classification (AIC 664 vs 684, and 137 vs 151 for OS and cumulative incidence of AML evolution, respectively). According to the genetic score, 41% of patients had a shift toward a higher risk category compared with the cytogenetic classification.

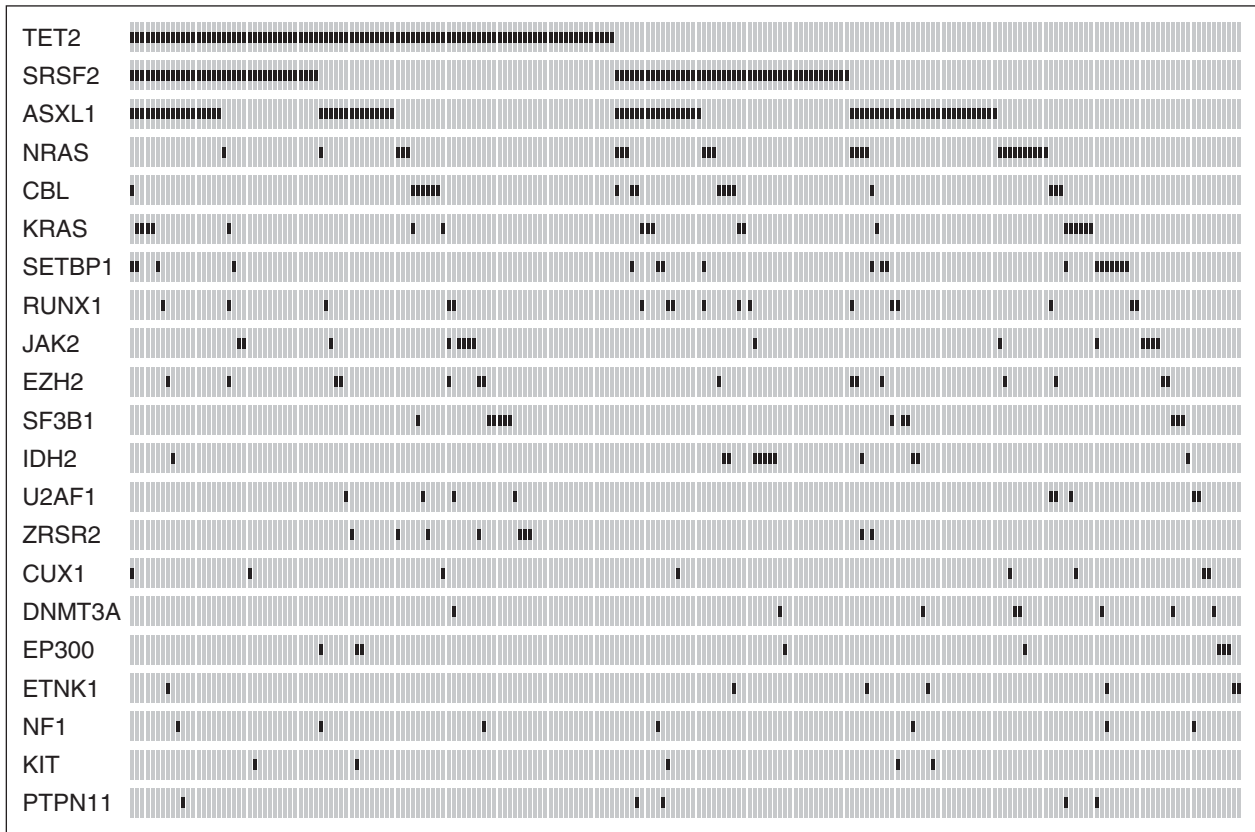


Figure 1. Mutation patterns observed in patients with CMML in the learning cohort. The plot represents the distribution of somatic lesions in genes mutated in $\geq 1\%$ of patients. Each column represents an individual patient sample.

Identification of hematologic prognostic variables and selection of cutoff values

Based on the observation that genetic parameters could not completely explain phenotypic variability, we then performed univariable and multivariable regression analyses including new genetic risk categories, clinical, and hematologic variables. In univariable analysis, the following phenotypic variables significantly influenced OS: hemoglobin value (HR = 0.8, $P < .001$), RBC transfusion dependency (HR = 2.8, $P < .001$), WBC count (HR = 1.02, $P < .001$), and BM blasts (HR = 1.11, $P < .001$). We also tested the value of sex-specific hemoglobin thresholds (ie, hemoglobin levels < 9 g/dL in males and < 8 g/dL in females) as an indicator of RBC transfusion dependency, as previously demonstrated in myelodysplastic syndromes (MDS).³² A high concordance was observed between sex-specific hemoglobin thresholds and RBC transfusion dependency (κ coefficient of concordance = .94, $P < .001$), and a comparable effect of the 2 parameters was found in univariable analysis (sex-specific hemoglobin thresholds: HR = 2.6, $P < .001$).

In order to identify the optimal cutoff values for the continuous variable to be integrated in the prognostic model, we calculated AIC of discrete values of BM blasts and WBC. The threshold of BM blast that was associated with the lowest AIC value was 7% (supplemental Figure 6A). Based on this observation and in order to have an easily clinically applicable discrete value, we adopted a threshold of 5% BM blasts (HR = 2.18, $P < .001$). The value of WBC count associated with the lowest AIC was $11 \times 10^9/L$ (supplemental Figure 6B). However, after pondering the trade-off between goodness of fit, expected impact on the model (10 out of 214 patients or 4.7%, had a WBC count between 11 and $13 \times 10^9/L$), as well as the unaccountable analytical sources of variance in WBC count,

we decided to maintain the threshold of $13 \times 10^9/L$ (HR = 2.46, $P < .001$).

Development of a clinical/molecular prognostic scoring system in CMML

In multivariable analysis, the following variables retained independent prognostic value: genetic risk groups (HR = 1.7, $P < .001$), RBC transfusion dependency (HR = 2.5, $P < .001$), WBC count (HR = 2.3, $P = .001$), and BM blasts (less than or equal to or $> 5\%$) (HR = 1.6, $P = .04$) (supplemental Table 5). Comparable results were obtained when sex-specific hemoglobin thresholds were used instead of RBC transfusion dependency in multivariable analysis (sex-specific hemoglobin thresholds, HR = 2.6, $P < .001$; genetic risk groups, HR = 1.7, $P < .001$; WBC count, HR = 2.1, $P = .002$; and BM blasts, HR = 1.6, $P = .038$).

Based on regression coefficients, we developed a clinical/molecular CPSS (CPSS-Mol), and the following scores were assigned to each variable: 1 point to intermediate-1 genetic score risk category, WBC $\geq 13 \times 10^9/L$, BM blasts $\geq 5\%$, and RBC transfusion dependency; 2 points to intermediate-2 genetic score; and 3 points to high genetic score (Table 3). This model was able to identify 4 risk groups with significantly different OS (HR = 2.69, $P < .001$) and cumulative incidence of leukemic evolution (HR = 3.84, $P < .001$) (median survival not reached, 64, 37, and 18 months; 48-month cumulative incidence of AML evolution of 0%, 3%, 21%, and 48% for the low, intermediate-1, intermediate-2, and high-risk group, respectively) (Figure 3A-B; supplemental Table 6).

We also calculated the scoring system using sex-specific hemoglobin thresholds instead of RBC transfusion dependency (κ coefficient

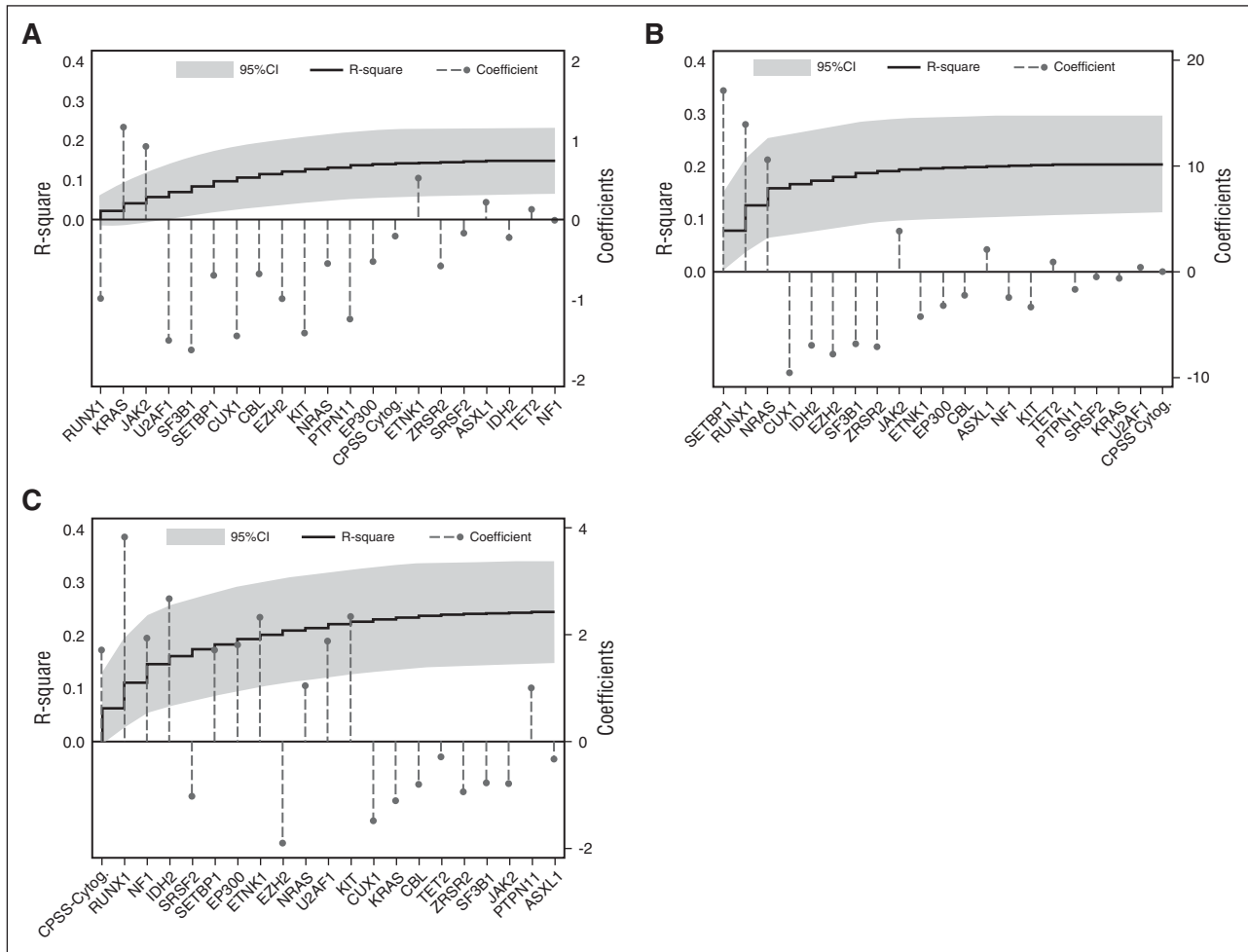


Figure 2. Multivariable models to predict hematologic variables from driver mutations. (A) Multivariable model to predict hemoglobin value from driver mutations. The step curve shows the cumulative proportion of variance (y-axis, left) in hemoglobin levels explained by each of the genetic variables. The gray shaded area represents the 95% confidence interval (CI) for this curve. Coefficient estimates for each gene in the model including all variables (y-axis, right) are shown as circles (coefficients >0 indicate positive correlation with hemoglobin levels, ie, the covariate is associated with higher hemoglobin levels; coefficients <0 indicate negative correlation with hemoglobin levels, ie, the covariate is associated with lower hemoglobin levels). (B) Multivariable model to predict WBC from driver mutations, as for (A). (C) Multivariable model to predict BM blast count from driver mutations, as for panels A-B.

of concordance = .95, $P < .001$), and comparable probability of OS and cumulative incidence of leukemic evolution were found (median survival of 100, 58, 37, and 18 months; 48-month cumulative incidence of AML evolution 0%, 9%, 16%, and 51% for the low, intermediate-1, intermediate-2, and high-risk group, respectively).

The AIC showed that the CPSS-Mol performed better than the original CPSS (AIC 630 vs 659, and 127 vs 145 for OS and cumulative incidence of AML evolution, respectively). According to CPSS-Mol,

48% of patients shifted to a higher risk group compared with the original CPSS.

Validation of the CPSS-Mol and comparison with previously defined prognostic scoring systems

Then, the prognostic value of the CPSS-Mol was tested in an independent validation cohort consisting of 260 CMML patients. When

Table 2. Variables and prognostic score values of the CMML genetic score

CPSS cytogenetic risk group*		ASXL1	NRAS	RUNX1	SETBP1
Variable score					
0	Low	Unmutated	Unmutated	Unmutated	Unmutated
1	Intermediate	Mutated	Mutated	—	Mutated
2	High	—	—	Mutated	—
Genetic risk group	Score				
Low	0				
Intermediate-1	1				
Intermediate-2	2				
High	≥3				

—, Not applicable.

*Cytogenetic risk groups are defined according to Such et al¹³: low, normal, and isolated -Y; intermediate, other abnormalities; and high, trisomy 8, complex karyotype (≥3 abnormalities), and abnormalities of chromosome 7.

Table 3. Variables and prognostic score values of the CPSS-Mol

	Genetic risk group*	BM blasts	WBC count	RBC transfusion dependency†
Variable score				
0	Low	<5%	<13 × 10 ⁹ /L	No
1	Intermediate-1	≥5%	≥13 × 10 ⁹ /L	Yes
2	Intermediate-2	—	—	—
3	High	—	—	—
CPSS-Mol risk group	Score			
Low	0			
Intermediate-1	1			
Intermediate-2	2-3			
High	≥4			

—, Not applicable.

*Genetic risk groups are defined as reported in Table 2.

†RBC transfusion dependency is defined according to Malcovati et al.³¹ and Such et al.¹³

comparing clinical and hematologic features of the learning and validation cohorts, a significant difference was found in the frequency of WHO categories, with CMML-2 being more frequent in the validation

cohort (38/214 vs 74/260 in the learning and validation cohort, respectively, *P* = .007), and accordingly of CPSS categories (*P* = .012). No significant difference was noticed between learning and validation

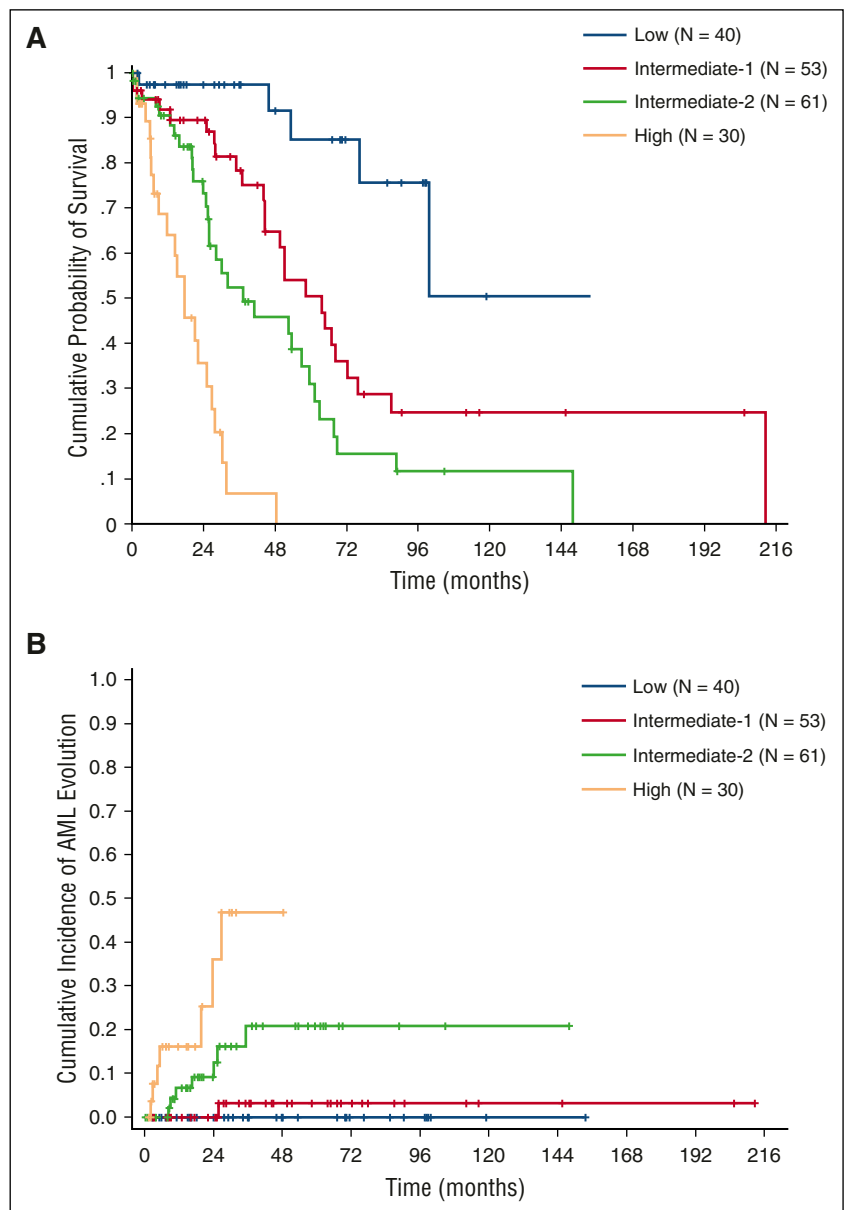


Figure 3. OS and cumulative incidence of leukemic evolution according to the CPSS-Mol in the learning cohort. (A) OS and (B) cumulative incidence of evolution into AML of patients classified into CPSS-Mol risk groups. The number of patients (N) in each category is reported: low risk group accounted for 22% of patients, intermediate-1 for 29%, and intermediate-2 and high-risk groups for 33% and 16% of patients, respectively.

cohorts in the prevalence of *ASXL1* (37% vs 40%, $P = \text{ns}$), *NRAS* (11.7% vs 13.4%, $P = \text{ns}$), and *SETBP1* mutations (8.9% vs 9.2%, $P = \text{ns}$), whereas a significantly higher prevalence of *RUNX1* mutations was observed in the validation cohort (20.7% vs 7.9%, $P < .001$).

Because the information of RBC transfusion-dependency was incomplete in the validation cohort, we adopted sex-specific hemoglobin thresholds to calculate the CPSS-Mol. The CPSS-Mol was able to identify 4 risk groups having a significantly different OS, with median survival time not reached, and 68, 30, and 17 months in the low, intermediate-1, intermediate-2, and high-risk group, respectively ($P < .001$). The 4 risk groups also showed a significantly different cumulative incidence of leukemic evolution, with a 48-month cumulative incidence of AML evolution of 0%, 8%, 24%, and 52% for the low, intermediate-1, intermediate-2, and high-risk group, respectively ($P < .001$) (Figure 4A-B; supplemental Table 6).

No significant difference was noticed in OS and cumulative incidence of AML evolution of CPSS-Mol risk groups between the learning and validation cohort, either when the CPSS-Mol was calculated in the learning cohort using RBC transfusion-dependency (P values ranging from .13 to .77, and from .22 to .99 for OS and cumulative incidence of AML evolution, respectively) or sex-specific hemoglobin thresholds (P values ranging from .37 to .78, and from .22 to .89 for OS and cumulative incidence of AML evolution, respectively).

We then calculated previously defined prognostic scoring systems including mutation status, according to Itzykson et al²⁰ and Patnaik et al,²⁶ and compared them with the CPSS-Mol in the learning cohort. The Groupe Francophone des Myelodysplasies (GFM) scoring system was based on the combination of age, hemoglobin, WBC and platelet values, and *ASXL1* mutation,²⁰ and was able to identify 3 risk groups with significantly different OS (HR = 2.16, $P < .001$) (median survival 64, 37, and 15 months, in the low, intermediate, and high-risk group, respectively) (supplemental Figure 7A). The Mayo Molecular model (MMM) was based on the combination of hemoglobin, absolute monocyte, circulating immature myeloid cells and platelet values, and *ASXL1* mutation,²⁶ and identified 4 risk groups with a median survival of 68, 50, 51, and 12 months, respectively (HR = 1.61, $P < .001$) (supplemental Figure 7B). The AIC and the Harrell's C concordance index showed that the CPSS-Mol performed better than either the GFM or the MMM (AIC: 630 vs 880 and 832; C-index: 0.73 vs 0.62 and 0.66 for the GFM and the MMM, respectively).

Discussion

In this study, we performed a comprehensive screening of somatic mutation on a large cohort of well-annotated patients with CMML to identify mutation patterns that may affect disease phenotype and clinical outcome. We observed significant correlations between genotype and disease phenotype. However, regression models showed that driver mutations are able to explain a limited fraction of the total variation of the most relevant clinical variables. Then, we refined cutoff values of the clinical parameters with significant prognostic value, and developed a clinical/molecular prognostic scoring system that was able to identify 4 risk groups with significantly different survival and risk of leukemic progression. The prognostic value of the scoring system was confirmed in an independent validation cohort.

CMML is characterized by an extremely heterogeneous disease phenotype, ranging from entities with predominant myelodysplastic features to forms with prevailing myeloproliferation.³⁷ Previous studies reported associations between somatic mutations and disease phenotype.^{20,24,38} We recognized mutation patterns significantly

associated with dysplastic or proliferative features, suggesting that these 2 disease subtypes may have at least partly different genetic background. However, we also found that recurrent somatic mutations were able to explain only a fraction of the total variation of the most relevant clinical variables, including hemoglobin, WBC count, and BM blasts. These data suggest that as-yet-unknown mutations or additional factors, including genetic background, clonal hierarchy, or distinct hematopoietic stem/progenitor cells primarily involved by mutation, may account for this unexplained tumor heterogeneity. These results also suggest that, at the current stage, the integration of genetic and clinical variables is expected to provide the maximal information for clinical decision-making.

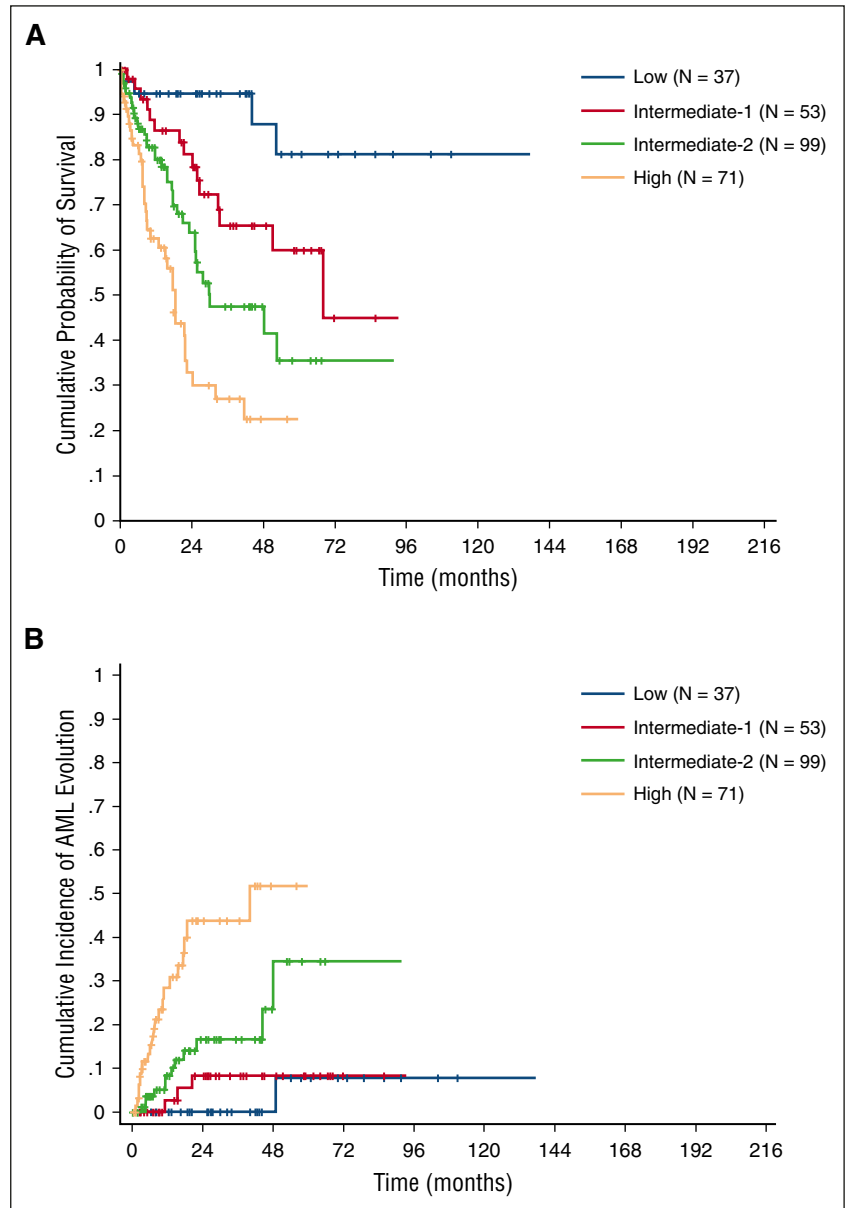
Several prognostic factors have been identified in CMML and combined into scoring systems, including clinical and hematologic variables.⁸⁻¹¹ Recently, we developed a CPSS,¹³ based on a comprehensive assessment of the prognostic value of cytogenetic abnormalities.¹⁴ Although the integration of cytogenetic abnormalities resulted in a better risk stratification, only a minority of CMML patients carry recurrent karyotypic abnormalities. Several genes were reported to have prognostic value in CMML.^{20,22,26,28,38,39} In 2 recent studies, the integration of *ASXL1* mutation status with clinical parameters allowed to develop prognostic scoring systems that appeared to be more discriminative than those based solely on clinical parameters.^{20,26}

We first explored the prognostic impact of genetic lesions, and confirmed the prognostic value of *ASXL1* mutations. In addition, we also found that somatic mutations in *RUNX1*, *NRAS*, and *SETBP1*, previously reported to have a prognostic value in MDS or MDS/MPN,^{28,40,41} had an additive independent prognostic value to CPSS cytogenetic risk groups. Based on multivariable regression coefficients, we defined a genetic score that was able to identify 4 risk groups, and was shown to perform better than the original CPSS cytogenetic risk classification. It must be acknowledged that our sample size did not allow to assess possible interactions between co-occurring mutations, which may potentially influence the resulting compounded prognostic effect.

We then investigated which clinical parameters were of additive value to this genetic classifier, and found that RBC transfusion-dependency, and WBC and BM blast counts retained an independent prognostic value. We also tested the value of sex-specific hemoglobin thresholds as an indicator of RBC transfusion-dependency, as previously defined in MDS,³² and obtained highly concordant results. Recent reports suggested that a threshold of 5% BM blasts was associated with a higher prognostic discrimination than the value of 10% introduced by the 2001 WHO classification.^{42,43} In order to identify the optimal cutoff value for continuous variable to be integrated in the prognostic model, we calculated the AIC of discrete values of BM blasts and WBC. We found that the value of 7% BM blasts was associated with highest goodness of fit. Therefore, consistent with the above-mentioned recent findings,^{42,43} we adopted a cutoff of 5% BM blasts. The WBC values associated with the best goodness of fit were between 11 and $12 \times 10^9/\text{L}$. However, it must be acknowledged that the identification of a clear WBC count value may be affected by several factors, including unaccountable analytical variation resulting in different upper normal values across laboratories.⁴⁴ In addition, a change in WBC threshold from 13 to $11 \times 10^9/\text{L}$ would affect risk assessment of <5% of patients included in this study. Therefore, although the criterion of $13 \times 10^9/\text{L}$ introduced by the FAB classification and confirmed by the recently updated WHO classification was associated with a slightly higher AIC value, in the development of the prognostic model we decided to maintain this definition, and devote to larger international initiatives the identification of the most appropriate WBC cutoff.

Based on the coefficients of the multivariable regression model, we defined a CPSS-Mol, which integrated the genetic score with RBC

Figure 4. OS and cumulative incidence of leukemic evolution according to the CPSS-Mol in the validation cohort. (A) OS and (B) cumulative incidence of evolution into AML of patients classified into CPSS-Mol risk groups. The number of patients (N) in each category is reported: low risk group accounted for 14% of patients, intermediate-1 for 20%, and intermediate-2 and high-risk groups for 38% and 27% of patients, respectively.



transfusion requirement, WBC and BM blast counts, and was able to identify 4 risk groups with significantly different survival and risk of leukemic evolution. As previously observed for the original CPSS, comparable results were obtained when the scoring system was calculated using hemoglobin thresholds instead of RBC transfusion dependency.¹³ This alternative model could be useful, especially in those patients who deserve urgent therapy and in whom RBC transfusion dependency is still not established.

The prognostic value of the CPSS-Mol was then tested in an independent cohort of patients with CMML. Because the information of RBC transfusion dependency was incomplete in the validation cohort, we adopted sex-specific hemoglobin thresholds. Although this validation cohort differed from the learning cohort in terms of morphologic subtypes and mutation pattern, no significant difference was noticed in OS and risk of AML evolution of CPSS-Mol risk groups between the learning and validation cohorts.

This prognostic score may be potentially useful for clinical decision-making in CMML. First, the CPSS-Mol was able to identify 2

categories with no or very low risk of disease progression, whose outcome was mainly affected by nonleukemic death related to BM failure. In addition, according to the CPSS-Mol, about half of patients had an estimate of their individual risk higher than that obtained with the original CPSS, and about one-third of patients with low or intermediate-1 CPSS risk were classified as intermediate-2 or high risk according to CPSS-Mol. In particular, CPSS-Mol was able to capture about 40% of patients with CMML-0 or -1 and a similar proportion of cases with normal karyotype as having an intermediate-2 or high risk based on their mutation profile. In addition, the CPSS-Mol performed better than previously defined *ASXL1* mutation-inclusive prognostic models,^{20,26} albeit it must be acknowledged that a comparison of goodness of fit within the data set on which the new model was developed is intrinsically biased.

In conclusion, although current understanding of the pathophysiology of CMML does not allow to move toward an entirely molecular-based classification and risk assessment, the results of this study suggest that the integration of somatic mutations significantly

improve prognostic stratification of patients with this neoplasm, and may provide a basis for improving clinical decision making.

Acknowledgments

This study was supported by grants from the Associazione Italiana per la Ricerca sul Cancro, Fondo per gli investimenti della ricerca di base (project no. RBAP11CZLK), and Ministero dell'Istruzione, dell'Università e della Ricerca (PRIN 2010-2011) (M.C.); the Associazione Italiana per la Ricerca sul Cancro (IG 15356) (L.M.); the Instituto de Salud Carlos III (grants PI12/01047, PIE13/00046, PT13/0010/0026 [Biobanco La Fe], and PI14/01649, and the "Consellería de Educación, Cultura y Deporte" of the Valencian Community (PROMETEOII/2015/008) (G.F.S.).

References

- Vardiman JW, Brunning R, Arber DA, et al. Introduction and overview of the classification of the myeloid neoplasms. In: Swerdlow SHCE, Harris NL, Jaffe ES, et al, eds. WHO Classification of Tumors Of Haematopoietic and Lymphoid Tissues. Lyon, France: IARC; 2008:18-30.
- Cazzola M, Malcovati L, Invernizzi R. Myelodysplastic/myeloproliferative neoplasms. *Hematology Am Soc Hematol Educ Program*. 2011;2011:264-272.
- Bennett JM, Catovsky D, Daniel MT, et al; Proposals by the French-American-British Cooperative Leukaemia Group. The chronic myeloid leukaemias: guidelines for distinguishing chronic granulocytic, atypical chronic myeloid, and chronic myelomonocytic leukaemia. *Br J Haematol*. 1994;87(4):746-754.
- Vardiman J, Pierre RV, Bain BJ, et al. Chronic myelomonocytic leukaemia. In: Swerdlow SHCE, Harris NL, Jaffe ES, et al, eds. WHO Classification of Tumors Of Haematopoietic and Lymphoid Tissues. Lyon, France: IARC; 2008:49-52.
- Arber DA, Orazi A, Hasserjian R, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*. 2016;127(20):2391-2405.
- Germing U, Strupp C, Knipp S, et al. Chronic myelomonocytic leukemia in the light of the WHO proposals. *Haematologica*. 2007;92(7):974-977.
- Aul C, Gattermann N, Heyll A, Germing U, Derigs G, Schneider W. Primary myelodysplastic syndromes: analysis of prognostic factors in 235 patients and proposals for an improved scoring system. *Leukemia*. 1992;6(1):52-59.
- Worsley A, Oscier DG, Stevens J, et al. Prognostic features of chronic myelomonocytic leukaemia: a modified Bournemouth score gives the best prediction of survival. *Br J Haematol*. 1988;68(1):17-21.
- Onida F, Kantarjian HM, Smith TL, et al. Prognostic factors and scoring systems in chronic myelomonocytic leukemia: a retrospective analysis of 213 patients. *Blood*. 2002;99(3):840-849.
- Greenberg P, Cox C, LeBeau MM, et al. International scoring system for evaluating prognosis in myelodysplastic syndromes. *Blood*. 1997;89(6):2079-2088.
- Greenberg PL, Tuechler H, Schanz J, et al. Revised international prognostic scoring system for myelodysplastic syndromes. *Blood*. 2012;120(12):2454-2465.
- Patnaik MM, Padron E, LaBorde RR, et al. Mayo prognostic model for WHO-defined chronic myelomonocytic leukemia: ASXL1 and spliceosome component mutations and outcomes [published correction appears in *Leukemia*. 2013;27(10):2112]. *Leukemia*. 2013;27(7):1504-1510.
- Such E, Germing U, Malcovati L, et al. Development and validation of a prognostic scoring system for patients with chronic myelomonocytic leukemia. *Blood*. 2013;121(15):3005-3015.
- Such E, Cervera J, Costa D, et al. Cytogenetic risk stratification in chronic myelomonocytic leukemia. *Haematologica*. 2011;96(3):375-383.
- Kohlmann A, Grossmann V, Klein HU, et al. Next-generation sequencing technology reveals a characteristic pattern of molecular mutations in 72.8% of chronic myelomonocytic leukemia by detecting frequent alterations in TET2, CBL, RAS, and RUNX1. *J Clin Oncol*. 2010;28(24):3858-3865.
- Meggendorfer M, Roller A, Haferlach T, et al. SRSF2 mutations in 275 cases with chronic myelomonocytic leukemia (CMML). *Blood*. 2012;120(15):3080-3088.
- Papaemmanuil E, Gerstung M, Malcovati L, et al; Chronic Myeloid Disorders Working Group of the International Cancer Genome Consortium. Clinical and biological implications of driver mutations in myelodysplastic syndromes. *Blood*. 2013;122(22):3616-3627, quiz 3699.
- Piazza R, Valletta S, Winkelmann N, et al. Recurrent SETBP1 mutations in atypical chronic myeloid leukemia. *Nat Genet*. 2013;45(1):18-24.
- Gambacorti-Passerini CB, Donadoni C, Parmiani A, et al. Recurrent ETNK1 mutations in atypical chronic myeloid leukemia. *Blood*. 2015;125(3):499-503.
- Itzykson R, Kosmider O, Renneville A, et al. Prognostic score including gene mutations in chronic myelomonocytic leukemia. *J Clin Oncol*. 2013;31(19):2428-2436.
- Maxson JE, Gotlib J, Pollyea DA, et al. Oncogenic CSF3R mutations in chronic neutrophilic leukemia and atypical CML. *N Engl J Med*. 2013;368(19):1781-1790.
- Kosmider O, Gelsi-Boyer V, Ciudad M, et al; Groupe Francophone des Myélodysplasies. TET2 gene mutation is a frequent and adverse event in chronic myelomonocytic leukemia. *Haematologica*. 2009;94(12):1676-1681.
- Malcovati L, Papaemmanuil E, Ambaglio I, et al. Driver somatic mutations identify distinct disease entities within myeloid neoplasms with myelodysplasia. *Blood*. 2014;124(9):1513-1521.
- Itzykson R, Kosmider O, Renneville A, et al. Clonal architecture of chronic myelomonocytic leukemias. *Blood*. 2013;121(12):2186-2198.
- Laborde RR, Patnaik MM, Lasho TL, et al. SETBP1 mutations in 415 patients with primary myelofibrosis or chronic myelomonocytic leukemia: independent prognostic impact in CMML. *Leukemia*. 2013;27(10):2100-2102.
- Patnaik MM, Itzykson R, Lasho TL, et al. ASXL1 and SETBP1 mutations and their prognostic contribution in chronic myelomonocytic leukemia: a two-center study of 466 patients. *Leukemia*. 2014;28(11):2206-2212.
- Gelsi-Boyer V, Trouplin V, Roquain J, et al. ASXL1 mutation is associated with poor prognosis and acute transformation in chronic myelomonocytic leukaemia. *Br J Haematol*. 2010;151(4):365-375.
- Damm F, Itzykson R, Kosmider O, et al. SETBP1 mutations in 658 patients with myelodysplastic syndromes, chronic myelomonocytic leukemia and secondary acute myeloid leukemias. *Leukemia*. 2013;27(6):1401-1403.
- Goasguen JE, Bennett JM, Bain BJ, Vallespi T, Brunning R, Mufti GJ; International Working Group on Morphology of Myelodysplastic Syndrome. Morphological evaluation of monocytes and their precursors. *Haematologica*. 2009;94(7):994-997.
- Mufti GJ, Bennett JM, Goasguen J, et al; International Working Group on Morphology of Myelodysplastic Syndrome. Diagnosis and classification of myelodysplastic syndrome: International Working Group on Morphology of myelodysplastic syndrome (IWGM-MDS) consensus proposals for the definition and enumeration of myeloblasts and ring sideroblasts. *Haematologica*. 2008;93(11):1712-1717.
- Malcovati L, Germing U, Kuendgen A, et al. Time-dependent prognostic scoring system for predicting survival and leukemic evolution in myelodysplastic syndromes. *J Clin Oncol*. 2007;25(23):3503-3510.
- Malcovati L, Della Porta MG, Strupp C, et al. Impact of the degree of anemia on the outcome of patients with myelodysplastic syndrome and its integration into the WHO classification-based Prognostic Scoring System (WPSS). *Haematologica*. 2011;96(10):1433-1440.
- Malcovati L, Della Porta MG, Pietra D, et al. Molecular and clinical features of refractory anemia with ringed sideroblasts associated with marked thrombocytosis. *Blood*. 2009;114(17):3538-3545.

34. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2010;26(5):589-595.
35. DePristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet*. 2011;43(5):491-498.
36. Fine J, Gray R. A proportional hazards model for the subdistribution of a competing risk. *J Am Stat Assoc*. 1999;94(446):496-509.
37. Mughal TI, Cross NC, Padron E, et al. An International MDS/MPN Working Group's perspective and recommendations on molecular pathogenesis, diagnosis and clinical characterization of myelodysplastic/myeloproliferative neoplasms. *Haematologica*. 2015;100(9):1117-1130.
38. Cervera N, Itzykson R, Coppin E, et al. Gene mutations differently impact the prognosis of the myelodysplastic and myeloproliferative classes of chronic myelomonocytic leukemia. *Am J Hematol*. 2014;89(6):604-609.
39. Kosmider O, Gelsi-Boyer V, Slama L, et al. Mutations of IDH1 and IDH2 genes in early and accelerated phases of myelodysplastic syndromes and MDS/myeloproliferative neoplasms. *Leukemia*. 2010;24(5):1094-1096.
40. Bejar R, Stevenson K, Abdel-Wahab O, et al. Clinical effect of point mutations in myelodysplastic syndromes. *N Engl J Med*. 2011;364(26):2496-2506.
41. Murphy DM, Bejar R, Stevenson K, et al. NRAS mutations with low allele burden have independent prognostic significance for patients with lower risk myelodysplastic syndromes. *Leukemia*. 2013;27(10):2077-2081.
42. Schuler E, Schroeder M, Neukirchen J, et al. Refined medullary blast and white blood cell count based classification of chronic myelomonocytic leukemias. *Leuk Res*. 2014;38(12):1413-1419.
43. Padron E, Garcia-Manero G, Patnaik MM, et al. An international data set for CMML validates prognostic scoring systems and demonstrates a need for novel prognostication strategies. *Blood Cancer J*. 2015;5:e333.
44. Savona MR, Malcovati L, Komrokji R, et al; MDS/MPN International Working Group. An international consortium proposal of uniform response criteria for myelodysplastic/myeloproliferative neoplasms (MDS/MPN) in adults. *Blood*. 2015;125(12):1857-1865.