

Integrated Clinical, Cytogenetic and Molecular Profiling in Chronic Myelomonocytic Leukemia: Genotype–Phenotype Associations and Risk-Model Discordance in an Indian Cohort.

Alpeshkumar Bipinbhai Kapadia

Christian Medical College, Vellore

Madhavi Maddali

Christian Medical College, Vellore

Phaneendra Datari

Christian Medical College, Vellore

Arun Kumar Arunachalam

Christian Medical College, Vellore

Sathya Mani

Christian Medical College, Vellore

Bexy Bensega

Christian Medical College, Vellore

Uday Prakash Kulkarni

Christian Medical College, Vellore

Anu Korula

Christian Medical College, Vellore

Biju George

Christian Medical College, Vellore

Aby Abraham

Christian Medical College, Vellore

Vikram Mathews

Christian Medical College, Vellore

Poonkuzhali Balasubramanian

bpoonkuzhali@cmcvellore.ac.in

Christian Medical College, Vellore

Keywords: chronic myelomonocytic leukemia, molecular profiling, cytogenetics, prognostic models, risk stratification

Posted Date: September 21st, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10970023/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

Chronic myelomonocytic leukemia (CMML) is a biologically heterogeneous myeloid neoplasm. Cytogenetic and molecular alterations are increasingly used for disease characterization and risk assessment. Indian data integrating molecular findings with clinical parameters and prognostic models remain limited.

Methods

We retrospectively analyzed 29 patients with newly diagnosed CMML evaluated between May 2021 and December 2025. Baseline clinical, cytogenetic, and molecular data were retrieved at the time of diagnosis. Mutations were characterized by gene frequency, functional pathway category, co-mutation pattern, and genotype–phenotype association across CMML subtypes. Baseline risk assignment and risk-class redistribution were evaluated using conventional and molecularly integrated CMML prognostic models.

Results

Median age was 65 years (IQR, 57–74 years), and 72.4% of patients were male. CMML-1 accounted for 75.9% of cases, and 62.1% had MP-CMML. Among cytogenetically evaluable cases, a normal karyotype was observed in 71.4%. The most commonly mutated genes were *TET2* (58.6%), *SRSF2* (44.8%), *ASXL1* (34.5%), *RUNX1* (27.6%), *KRAS* (24.1%), and *NRAS* (20.7%). Epigenetic regulator and signaling pathway alterations were observed in 89.7% and 69.0% of cases, respectively. In exploratory co-mutation analysis, *TET2*–*SRSF2* was the most frequently observed co-mutation pair, and *KRAS* mutations clustered in MP-CMML in exploratory analysis. Baseline risk assignment varied considerably across prognostic models.

Conclusion

Integrated clinicogenomic profiling delineates substantial biological heterogeneity in CMML and model-dependent variation in baseline CMML risk assignment in this Indian cohort.

Introduction

Chronic myelomonocytic leukemia (CMML) is a clonal hematopoietic stem and progenitor cell disorder with overlapping myelodysplastic and myeloproliferative features, characterized by persistent monocytosis and an increased risk of transformation to acute myeloid leukemia [1–3]. According to the fifth Edition of World Health Organization of hematolymphoid tumors (WHO HAEM5), CMML is defined by sustained peripheral blood (PB) monocytosis, with both an absolute monocyte count of $\geq 0.5 \times 10^9/L$

and monocytes comprising $\geq 10\%$ of leukocytes, exclusion of other well-established and molecularly defined myeloproliferative neoplasms, PB and/or bone marrow (BM) blast equivalents $< 20\%$, and supportive evidence of clonality and/or dysplasia [1]. WHO HAEM5 subclassified CMML by blast percentage into CMML-1 and CMML-2, and removed CMML-0 as a subcategory used in previous classifications [4]. CMML-1 includes cases with $< 5\%$ blasts in PB and $< 10\%$ blasts in BM, whereas CMML-2 includes cases with 5–19% blasts in PB, 10–19% blasts in BM, and/or the presence of Auer rods [1]. In addition, CMML is clinically categorized into myelodysplastic CMML (MD-CMML) and myeloproliferative CMML (MP-CMML) using a total leukocyte count threshold of $13 \times 10^9/\text{L}$, reflecting the important biological and clinical distinction between cytopenic and proliferative disease phenotypes [1, 2].

With the availability of high-throughput next-generation sequencing technologies, recurrent somatic mutations have emerged as central determinants of CMML biology [5–13]. Large sequencing studies have shown that CMML is enriched for mutations in genes regulating epigenetic modification, RNA splicing, chromatin remodeling, transcriptional regulation, and RAS/MAPK signaling [7, 8]. Alterations in *TET2*, *SRSF2*, *ASXL1*, and *RUNX1* define a major component of the CMML mutational landscape, with *TET2-SRSF2* co-mutation representing a recurrent early cooperative event, while signaling pathway lesions involving *KRAS*, *NRAS*, *CBL*, *NF1*, and *PTPN11* are associated with proliferative phenotype, clonal evolution, and clinical progression [7, 8, 11–16].

Risk assessment in CMML has evolved from clinical and cytogenetic models to molecularly integrated systems [17–19]. Earlier prognostic models, including the MD Anderson Prognostic Score (MDAPS) and Mayo Prognostic Model (MPM), relied mainly on hematologic parameters [17, 20, 21]. The CMML-Specific Prognostic Scoring System (CPSS) subsequently added cytogenetic risk to clinical variables [18]. With growing recognition of the diagnostic and prognostic relevance of molecular alterations, models such as the Groupe Français des Myélodysplasies (GFM), Mayo Molecular Model (MMM), CPSS-Mol, BLAST-Mol, and international CMML Prognostic Scoring System (iCPSS) have expanded risk assessment by incorporating molecular information [18, 19, 22, 23].

Despite growing recognition of the prognostic significance of recurrent gene mutations in CMML, Indian data integrating clinical, cytogenetic, and molecular features remain scarce [24]. In particular, the extent to which conventional clinical, cytogenetic, molecular, and integrated prognostic systems categorize baseline risk in Indian CMML patients has not been systematically evaluated. Therefore, we aimed to describe the clinical, cytogenetic, and molecular landscape of CMML in an Indian tertiary-care cohort, and to explore co-mutation patterns, genotype-phenotype relationships, and differences in baseline risk assignment across multiple conventional and molecularly integrated CMML prognostic models.

Materials and Methods

Study Design and Patient Cohort

This retrospective observational study included newly diagnosed CMML patients evaluated at our tertiary care center between May 2021 and December 2025. Cases were identified through the diagnostic molecular testing database and included if baseline targeted next-generation sequencing data were available, and conventional cytogenetic analysis had been performed or attempted at diagnosis. Baseline clinical, hematologic, cytogenetic, and molecular data were retrieved from electronic medical records and laboratory databases at diagnosis or the closest available time point. All cases were diagnosed as CMML following integrated review of clinical, hematologic, cytogenetic, and molecular findings, using WHO HAEM5 as the primary diagnostic framework. Because the cohort was derived from patients suspected of having CMML who underwent molecular testing, the total morphology-based CMML denominator could not be reliably determined. The study was approved by the Institutional Ethics Committee.

Cytogenetic analysis

Conventional G-banded cytogenetic analysis was performed on diagnostic bone marrow samples using standard procedures [25]. CMML-specific cytogenetic risk was assigned per Such et al. classification [26]. Briefly, low risk included a normal karyotype and Y chromosome loss; high risk included trisomy 8, chromosome 7 abnormalities, and complex karyotype; and all others were classified intermediate risk. Cases with incomplete or unsuccessful cytogenetic studies were retained for clinical and molecular analyses but excluded from cytogenetic-risk assignment and from prognostic models requiring evaluable cytogenetic data.

Molecular analysis

Targeted next-generation sequencing was performed on diagnostic bone marrow or peripheral blood samples using a 57-gene myeloid panel (**Supplementary Table S1**).

Libraries were prepared from 50–250 ng of genomic DNA using a UMI-based targeted hybrid-capture assay with a custom myeloid gene capture kit and sequenced on the Illumina HiSeq X or NovaSeq 6000 platforms. Further assay and bioinformatics details are provided in the **Supplementary Methods**. Clinically relevant pathogenic or likely pathogenic variants were included using a minimum variant allele frequency threshold of 5%. Variant interpretation was performed according to AMP/ASCO/CAP and ClinGen/CGC/VICC recommendations [27, 28]. Matched germline control samples were not available; likely germline polymorphisms were assessed using curated population and disease databases, although germline origin could not be fully excluded. Copy-number alterations (CNAs) and structural variants, including gene fusions, were not evaluated, as the DNA-based assay was not validated for CNAs and lacked fusion-detection capability. Mutation status was recorded as present or absent, and the number of mutations per patient was calculated.

TET2 alterations were categorized as single-hit or putative multi-hit. Putative multi-hit status was assigned when multiple pathogenic *TET2* variants were identified or when a single variant had a VAF $\geq$ 50% [12]. Because copy-number analysis, phasing, tumor purity assessment, and matched germline testing were unavailable, the true biallelic status could not be confirmed.

Functional Pathway, Co-Mutation, and Genotype–Phenotype Analysis

Mutated genes were grouped into functional categories, including epigenetic regulation, signaling/RAS-MAPK pathway, spliceosome, transcriptional regulation, cohesin complex, and nuclear-cytoplasmic transport. Pathway frequencies were calculated per patient and were not mutually exclusive. Co-mutation analysis was performed on recurrently mutated genes, defined as those mutated in $\geq$ 10% of cases or $\geq$ 3 patients. Genotype-phenotype correlations were evaluated across CMML subgroups, including CMML-1 versus CMML-2 and MD-CMML versus MP-CMML. Due to the small sample size, the co-mutation and genotype-phenotype analyses were exploratory in nature.

Prognostic Model Assignment and Risk-Class Redistribution

Available CMML prognostic models were grouped by the types of variables incorporated. Clinical-only models included MDAPS, MPM, and BLAST. The clinical-cytogenetic model was CPSS. Clinical-molecular models included GFM and MMM. Integrated models included CPSS-Mol, BLAST-Mol, and iCPSS. Risk categories followed the original definitions, and cytogenetic models included only evaluable cases. Baseline risk was summarized, and risk-class redistribution was assessed through paired comparisons such as MPM to MMM, BLAST to BLAST-Mol, and CPSS to CPSS-Mol/iCPSS.

Statistical Analysis

Continuous variables were summarized as median and interquartile range (IQR), and categorical variables as counts and percentages. Continuous variables were compared using the Mann-Whitney U test, whereas categorical variables were compared using Fisher's exact test. All statistical tests were two-sided, and P values < 0.05 were considered statistically significant. The phi coefficient (ϕ) was used to assess co-occurrence or mutual exclusivity between mutation pairs, with Benjamini–Hochberg false discovery rate (FDR) correction ($q < 0.10$) for significance. Data processing and visualization, including oncoplots, mutation distribution plots, and risk-category migration plots, were performed using Python-based libraries [29–31].

Results

Baseline Clinical and Laboratory Characteristics

Twenty-nine patients with newly diagnosed CMML were included, and their baseline characteristics are summarized in Table 1. The median age at diagnosis was 65 years (IQR, 57–74 years), with a marked male predominance (72.4%). Most patients were cytopenic, with ~ 76% having hemoglobin below 10 g/dL and nearly 70% having thrombocytopenia ($< 100 \times 10^9/L$). Leukocytosis was common, and 62.1% were classified as myeloproliferative CMML (MP-CMML) and 37.9% as myelodysplastic CMML (MD-CMML). Peripheral blood blasts and circulating immature myeloid cells were identified in 34.5% and 75.9% of patients, respectively. Nearly half of the cohort was transfusion-dependent at the time of diagnosis. Based on the WHO blast-based classification, the majority of cases were CMML-1 (75.9%), with CMML-2 accounting for 24.1%.

Table 1
Baseline Clinical and hematologic characteristics at diagnosis

Characteristic	All patients (n = 29)
Age in years, median (IQR)	65 (57–74)
≤ 60 years, n (%)	9 (31.0)
61–70 years, n (%)	11 (37.9)
> 70 years, n (%)	9 (31.0)
Sex, n (%)	
Male	21 (72.4)
Female	8 (27.6)
WHO subtype, n (%)	
CMML-1	22 (75.9)
CMML-2	7 (24.1)
Hemoglobin, g/dL, median (IQR)	8.7 (7.9–9.8)
< 10 g/dL, n (%)	22 (75.9)
≥ 10 g/dL, n (%)	7 (24.1)
Total Leukocyte count, ×10⁹/L, median (IQR)	20.6 (8.7–31.0)
< 13×10 ⁹ /L (MD-CMML), n (%)	11 (37.9)
≥ 13×10 ⁹ /L (MP-CMML), n (%)	18 (62.1)
Platelet count, ×10⁹/L, median (IQR)	71 (47–123)
< 100×10 ⁹ /L, n (%)	20 (69.0)
≥ 100×10 ⁹ /L, n (%)	9 (31.0)
Absolute monocyte count, ×10⁹/L, median (IQR)	3.4 (1.9–6.9)
< 10×10 ⁹ /L, n (%)	26 (89.7)
≥ 10×10 ⁹ /L, n (%)	3 (10.3)
Peripheral blood blasts, median (IQR)	0 (0–2)
Absent, n (%)	19 (65.5)
Present, n (%)	10 (34.5)
Bone marrow blasts, median (IQR)	4 (2–7)
< 5%, n (%)	15 (51.7)

Characteristic	All patients (n = 29)
Age in years, median (IQR)	65 (57–74)
5–9%, n (%)	9 (31.0)
≥ 10%, n (%)	5 (17.2)
Circulating immature myeloid cells	
Absent, n (%)	7 (24.1)
Present, n (%)	22 (75.9)
LDH, U/L*, median (IQR)	326 (250–456)
Transfusion dependence	
No, n (%)	15 (51.7)
Yes, n (%)	14 (48.3)
Chromosomal abnormalities, n (%)†	
Normal karyotype, n (%)	20 (71.4)
Trisomy 8, n (%)	4 (14.3)
Monosomy 7, n (%)	2 (7.1)
Minus Y, n (%)	1 (3.6)
inv(2)(p23q13), n (%)	1 (3.6)
CMML-specific cytogenetic risk classification†	
Low	21 (75.0)
Intermediate	1 (3.6)
High	6 (21.4)
Molecular Features	
NGS positive, n (%)	27 (93.1)
NGS negative, n (%)	2 (6.9)
Number of mutations per patient, median (IQR)	3 (3–5)

* LDH available in 26 patients.

†Chromosomal abnormalities and CMML-specific cytogenetic risk classification were evaluable in 28 cases; one patient was excluded because of an incomplete cytogenetic study.

Cytogenetic Landscape

Conventional cytogenetic analysis was performed in 29 patients, with one case showing an incomplete study (Table 1). Of 28 evaluable cases, a normal karyotype was observed in 20 patients (71.4%). The most common abnormalities were trisomy 8 (4 cases, 14.3%) and monosomy 7 (2 cases, 7.1%); loss of Y and inv(2)(p23q13) were each observed once (3.6%). By CMML-specific cytogenetic risk, 21 of the 28 cases were low risk (21, 75.0%), with 6 (21.4%) high risk and 1 (3.6%) intermediate.

Molecular Landscape and Mutation Spectrum

Targeted next-generation sequencing identified pathogenic or likely pathogenic alterations in 27 of 29 patients (93.1%), with a median of 3 mutations per patient. *TET2* was the most frequently mutated gene, detected in 17 patients (58.6%). Putative multi-hit *TET2* alterations were identified in 13 patients (44.8%), while 4 (13.8%) had single-hit alterations. Other recurrently mutated genes included *SRSF2* (n = 13, 44.8%), *ASXL1* (n = 10, 34.5%), *RUNX1* (n = 8, 27.6%), *KRAS* (n = 7, 24.1%), and *NRAS* (n = 6, 20.7%) (Fig. 1). Less frequent mutations included *CBL*, *NF1*, and *PTPN11*, each detected in 3 patients (10.3%), whereas other alterations were detected in only 1 or 2 patients (Fig. 1). Two patients harbored *NPM1* mutations, had persistent monocytosis, CMML-like morphology, and < 10% blasts, and were clinically diagnosed and managed as CMML. Because the WHO HAEM5 and ICC 2022 differ in their classification of *NPM1*-mutated myeloid neoplasms, these cases were retained for descriptive analysis but regarded as diagnostically ambiguous.

At the pathway level, alterations most frequently involved epigenetic regulators (89.7%), followed by signaling pathway genes (69.0%), spliceosome genes (48.3%), transcription factor genes (31.0%), cohesin complex genes (6.9%), and nuclear-cytoplasmic transport genes (6.9%) (Fig. 1, **lower panel**).

Pairwise mutational association analysis

Exploratory pairwise mutation association analysis revealed *TET2*–*SRSF2* as the most frequent co-mutation pair and the only association that remained significant after FDR correction (n = 12, ϕ = 0.60; Fisher's exact *P* = 0.0018, OR: 23.0, 95% CI: 2.3-1227.7). However, the odds-ratio estimate had a wide confidence interval, reflecting the small cohort size. No other pairs were significant after FDR correction. Positive but non-significant associations were observed for *TET2*–*RUNX1* (7 cases, ϕ = 0.40), *SRSF2*–*RUNX1* (6 cases, ϕ = 0.40), *TET2*–*KRAS* (6 cases, ϕ = 0.30), and *ASXL1*–*NRAS* (4 cases, ϕ = 0.30), whereas *TET2*–*ASXL1* showed a moderate negative association (4 cases, ϕ = -0.30). None of these secondary associations remained significant after FDR correction (**Supplementary Fig. S1**).

Genotype-phenotype correlations

Exploratory genotype–phenotype analyses were performed across blast-defined CMML categories and proliferative subtypes. Compared with CMML-1, CMML-2 cases had higher median leukocyte count [41.8 $\times 10^9$ /L vs. 18.5 $\times 10^9$ /L], absolute monocyte count [5.02 $\times 10^9$ /L vs. 3.17 $\times 10^9$ /L], and LDH level [415 U/L vs. 317.5 U/L]. CMML-2 cases also had a higher frequency of an abnormal karyotype [4/7, 57.1% vs. 4/21, 19.0%] and a CPSS high-risk cytogenetic assignment [3/7, 42.9% vs. 3/21, 14.3%]. Among recurrently mutated genes, *ASXL1* [4/7, 57.1% vs. 6/22, 27.3%], *RUNX1* [3/7, 42.9% vs. 5/22, 22.7%],

SRSF2 [4/7, 57.1% vs. 9/22, 40.9%], *KRAS* [2/7, 28.6% vs. 5/22, 22.7%], and *NRAS* [2/7, 28.6% vs. 4/22, 18.2%] were numerically more frequent in CMML-2, whereas *TET2* mutations were more frequent in CMML-1 [14/22, 63.6% vs. 3/7, 42.9%]. However, no clinical, cytogenetic, or recurrent molecular features differed significantly between the groups (Fig. 2).

In the MD-CMML versus MP-CMML comparison, MP-CMML cases had significantly higher median leukocyte [$28.5 \times 10^9/L$ vs. $7.8 \times 10^9/L$; $P < 0.001$], and higher absolute monocyte counts [$6.34 \times 10^9/L$ vs. $1.49 \times 10^9/L$; $P < 0.001$]. *KRAS* mutations were detected in 7 of 18 (38.9%) MP-CMML cases and in none of the 11 MD-CMML cases [$P = 0.025$, OR 15.0, 95% CI 0.76–294.5]. However, this nominal association did not remain significant after FDR correction (FDR-adjusted $P = 0.519$) and should be considered exploratory rather than definitive evidence of an association. No other recurrent mutation differed significantly between the subtypes (**Supplementary Fig. S2**).

Risk distribution and risk-class redistribution across CMML prognostic models

The assignment of risk categories varied substantially among the different prognostic models evaluated. The proportion of patients classified as high risk ranged from 14.3% with CPSS to 79.3% with MPM and BLAST. Models based solely on clinical data tended to group patients into higher-risk categories, while models incorporating molecular and/or cytogenetic variables produced more heterogeneous risk distributions. Specifically, CPSS-Mol classified 35.7% as high risk, BLAST-Mol classified 50.0% as high risk, and iCPSS assigned patients across the entire spectrum from very low to very high risk. Detailed model-specific distributions are provided in **Supplementary Table S2** and are summarized in Fig. 3.

Paired comparisons were performed to describe differences in baseline risk assignment between related clinical and molecularly integrated models. In the MPM to MMM comparison, no patient was classified as low risk by either model. Of the 6 MPM intermediate risk patients, 5 (83.3%) remained intermediate-1 and 1 (16.7%) was reassigned to intermediate-2 by MMM. The high-risk fraction decreased from 23 of 29 (79.3%) with MPM to 16 of 29 (55.2%) with MMM, with 7 of 23 (30.4%) MPM high-risk patients reassigned to intermediate-2. Similarly, among the 28 patients evaluable by both BLAST and BLAST-Mol, the low- and intermediate-risk cases remained in their corresponding categories, whereas 9 of 23 (39.1%) BLAST high-risk patients were reassigned to intermediate risk. As a result, the proportion of high-risk cases decreased from 23 of 28 (82.1%) with BLAST to 14 of 28 (50.0%) with BLAST-Mol (**Supplementary Fig. S3**).

Conversely, CPSS-Mol shifted the overall distribution toward higher risk categories relative to CPSS. Both CPSS low risk cases were reassigned to intermediate-risk categories. Among the 12 CPSS intermediate-1 cases, 7 (58.3%) were reassigned to intermediate-2 and 2 (16.7%) to high risk, while 4 of 10 (40.0%) CPSS intermediate-2 cases were reassigned to high risk. Consequently, the high-risk fraction increased from 4 of 28 (14.3%) with CPSS to 10 of 28 (35.7%) with CPSS-Mol. In contrast, iCPSS redistributed patients across a broader risk spectrum, with 9 of 28 (32.1%) classified as very low or low risk, 9 of 28

(32.1%) as intermediate risk, and 10 of 28 (35.7%) as high or very high risk. Thus, the direction and magnitude of baseline risk redistribution varied across models (Fig. 4).

Discussion

In this retrospective cohort of 29 newly diagnosed CMML patients, we integrated baseline clinical features, conventional cytogenetics, and targeted next-generation sequencing to characterize disease heterogeneity and compare baseline risk assignment between conventional and molecularly integrated prognostic models. The analysis extended beyond mutation-frequency reporting to include functional pathway involvement, exploratory co-mutation and genotype-phenotype patterns, and model-dependent risk redistribution. Since Indian data on CMML are scarce, this study provides a descriptive framework for understanding the biological heterogeneity and variability in baseline risk classification in the Indian tertiary-care setting.

The cohort had a median age of 65 years, male predominance, and a distribution of dysplastic and proliferative subtypes broadly consistent with previous studies [17, 18, 21, 32]. However, our cohort showed a high occurrence of adverse baseline features, including anemia, thrombocytopenia, circulating immature myeloid cells, transfusion dependence, CMML-2, and MP-CMML. This enrichment likely reflects tertiary-care referral and molecular-testing selection biases, as patients were identified through a diagnostic molecular database rather than an unselected, morphology-based CMML registry. Patients with aggressive, proliferative, atypical, or diagnostically challenging disease may have been more likely to undergo NGS and cytogenetic evaluation. Consequently, the prevalence of adverse clinical features and high-risk assignments may be overestimated and should not be interpreted as population-level estimates for CMML in India [21, 23, 33]. A previous Indian study reported MP-CMML in 40% of patients, compared with 62.1% in our cohort, suggesting that differences between studies are probably due to cohort composition rather than to true population-level biological variation [24].

Cytogenetic analysis showed a normal karyotype in 71.4% of patients, while recurrent abnormalities included trisomy 8 or monosomy 7, consistent with previous CMML studies [9, 26, 32]. Despite the high prevalence of adverse clinical features, most patients were assigned to the low risk cytogenetic category. This apparent divergence highlights the limited ability of cytogenetic findings alone to capture the full clinical and molecular heterogeneity of CMML and provides a rationale for examining how molecularly integrated models alter baseline risk assignment [17–19, 22].

The mutational profile recapitulated the known molecular landscape of CMML. Genes such as *TET2*, *SRSF2*, *ASXL1*, and *RUNX1* were among the most frequently altered, consistent with findings from large CMML studies [23, 33]. Notably, *TET2* mutations, particularly putative multi-hit mutations, are prominent, as they are characteristic of overt and oligomonocytic CMML [34–36]. The frequent *SRSF2* mutations and their co-mutation with *TET2* support the biological cooperation between epigenetic and spliceosomal alterations recognized in CMML pathogenesis, although the magnitude of association in this small cohort should be interpreted cautiously [5, 8, 16]. Additionally, RAS/MAPK pathway alterations, found in about 69% of patients, may have contributed to the predominance of MP-CMML in this cohort,

consistent with the established relationship between signaling mutations and MP-CMML biology [5, 7, 11, 19]. *KRAS* mutations clustered in MP-CMML, but the association did not remain significant after FDR correction. This observation is directionally consistent with the role of RAS/MAPK activation in proliferative CMML but should be regarded as hypothesis-generating, particularly given the small sample size.

Application of the prognostic models revealed substantial variability in baseline risk assignment across models. Clinical-only models, particularly MPM and BLAST, classified most cases as high risk, whereas CPSS distributed patients more broadly across intermediate-risk categories. These differences likely reflect the variables and thresholds included in each model [18, 21, 33]. Similarly, although both GFM and MMM incorporate hemoglobin levels and *ASXL1* mutation status, differences in the inclusion and weight assigned to other clinical variables—such as age, cytopenias, proliferative burden, monocytosis, and circulating immature myeloid cells—probably explain their differing risk distributions [22, 37]. Differences between the distributions reported in the original development cohorts and the present cohort should therefore be interpreted in the context of referral patterns, testing selection, eligibility criteria, and model structure, rather than as evidence of population-specific biological differences.

Molecularly integrated models also showed differences in the extent of baseline risk redistribution. Compared with their clinical counterparts, MMM and BLAST-Mol reclassified a subset of clinically high-risk patients as intermediate risk. Conversely, CPSS-Mol reclassified CPSS intermediate-risk cases to intermediate-2 or high risk, reflecting the combined contributions of clinical and cytogenetic variables, as well as mutations in *ASXL1*, *RUNX1*, *NRAS*, and *SETBP1* [19, 38]. iCPSS distributed patients over a wider risk range, probably due to its combined weighting of hematologic, cytogenetic, and molecular factors [23]. This finer stratification aligns with emerging molecular classifications that yield a broader categorical distribution, although it does not confer superior discrimination in the absence of outcome data. This approach is particularly relevant in modern CMML care, where molecular markers are increasingly complementing conventional risk categories through molecular-based disease classification [12, 23, 39].

The strengths of this study include the systematic integration of clinical, cytogenetic, and targeted NGS findings and the concurrent evaluation of functional pathways, co-mutation patterns, genotype–phenotype associations, and multiple prognostic models in an Indian CMML cohort. However, these findings should be interpreted with caution given the retrospective, single-center design, small sample size, limited power for subgroup and pairwise analyses, and potential referral and molecular testing selection biases. Furthermore, the absence of matched germline testing, mutation phasing, copy-number assessment, and tumor-purity estimation also prevented definitive confirmation of biallelic *TET2* involvement. The inclusion of two *NPM1*-mutated cases with monocytosis, CMML-like morphology, and < 10% blasts represents an additional classification limitation, as WHO HAEM5 and ICC 2022 differ in their categorization of *NPM1*-mutated myeloid neoplasms, placing these cases within an area of diagnostic uncertainty. These cases were retained to reflect real-world diagnostic practice, but their inclusion may have influenced the reported mutation frequencies and risk-model classifications. Finally,

because survival and AML-transformation outcomes were unavailable, the observed differences across prognostic systems should be interpreted as model-dependent variations in baseline risk assignment and do not establish the prognostic superiority of any individual model in this cohort.

In conclusion, this cohort recapitulated key features of CMML biology, including frequent alterations involving epigenetic regulation, spliceosome machinery, and RAS/MAPK signaling. The *TET2*–*SRSF2* co-mutation and clustering of *KRAS* mutations in MP-CMML represent biologically consistent but predominantly exploratory findings. Prognostic models showed substantial model-dependent variation in baseline risk assignments, with molecularly integrated systems exhibiting model-specific rather than uniformly upward redistribution. Larger multicenter studies incorporating survival and AML-transformation endpoints are needed to determine the clinical relevance of these observations in Indian patients with CMML.

Declarations

Acknowledgements:

The authors acknowledge the laboratory haematology and pathology teams involved in diagnostic work-up.

Authors' contributions

ABK and PB conceived and designed the study. ABK performed the analysis, interpreted the data, and drafted the manuscript. MM and PD contributed to data analysis and interpretation. SM and BB collected and analyzed the data. UPK, AK, BJ, AA, and VM contributed to patient management, data interpretation, and critical review of the manuscript. PB supervised the study and critically revised the manuscript. All authors reviewed and approved the final submitted version.

Ethics approval and consent to participate: The study was approved by the Institutional Review Board. The requirement for individual informed consent was waived because of the retrospective nature of the study and the use of de-identified clinical and laboratory data.

Data availability: The anonymized data supporting the findings of this study are included in the manuscript and its supplementary materials. Additional de-identified data may be made available by the corresponding author upon reasonable request and subject to institutional approval.

Funding: This study was not supported by any sponsor or funder.

Conflict of Interest: The authors declare no competing interests.

References

1. Khoury JD, Solary E, Abela O et al (2022) The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Myeloid and Histiocytic/Dendritic Neoplasms. *Leukemia* 36:1703–1719. <https://doi.org/10.1038/s41375-022-01613-1>
2. Patnaik MM, Tefferi A (2024) Chronic myelomonocytic leukemia: 2024 update on diagnosis, risk stratification and management. *Am J Hematol* 99:1142–1165. <https://doi.org/10.1002/ajh.27271>
3. Valent P, Orazi A, Savona MR et al (2019) Proposed diagnostic criteria for classical chronic myelomonocytic leukemia (CMML), CMML variants and pre-CMML conditions. *Haematologica* 104:1935–1949. <https://doi.org/10.3324/haematol.2019.222059>
4. Arber DA, Orazi A, Hasserjian R et al (2016) The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood* 127:2391–2405. <https://doi.org/10.1182/blood-2016-03-643544>
5. Itzykson R, Kosmider O, Renneville A et al (2013) Clonal architecture of chronic myelomonocytic leukemias. *Blood* 121:2186–2198. <https://doi.org/10.1182/blood-2012-06-440347>
6. Kar SA, Jankowska A, Makishima H et al (2013) Spliceosomal gene mutations are frequent events in the diverse mutational spectrum of chronic myelomonocytic leukemia but largely absent in juvenile myelomonocytic leukemia. *Haematologica* 98:107–113. <https://doi.org/10.3324/haematol.2012.064048>
7. Patel BJ, Przychodzen B, Thota S et al (2017) Genomic determinants of chronic myelomonocytic leukemia. *Leukemia* 31:2815–2823. <https://doi.org/10.1038/leu.2017.164>
8. Palomo L, Meggendorfer M, Hutter S et al (2020) Molecular landscape and clonal architecture of adult myelodysplastic/myeloproliferative neoplasms. *Blood* 136:1851–1862. <https://doi.org/10.1182/blood.2019004229>
9. Coltro G, Mangaonkar AA, Lasho TL et al (2020) Clinical, molecular, and prognostic correlates of number, type, and functional localization of TET2 mutations in chronic myelomonocytic leukemia (CMML)—a study of 1084 patients. *Leukemia* 34:1407–1421. <https://doi.org/10.1038/s41375-019-0690-7>
10. Han W, Zhou F, Wang Z et al (2022) Mutational landscape of chronic myelomonocytic leukemia and its potential clinical significance. *Int J Hematol* 115:21–32. <https://doi.org/10.1007/s12185-021-03210-x>
11. Montalban-Bravo G, Rodriguez-Sevilla JJ, Swanson DM et al (2024) Influence of co-mutational patterns in disease phenotype and clinical outcomes of chronic myelomonocytic leukemia. *Leukemia* 38:1178–1181. <https://doi.org/10.1038/s41375-024-02190-1>
12. Montalban-Bravo G, Wu C, Rodriguez-Sevilla JJ et al (2026) Incorporation of Genomic Determinants Improves Diagnostic Accuracy of Oligomonocytic Chronic Myelomonocytic Leukemia. *Blood Cancer Discovery* 7:414–430. <https://doi.org/10.1158/2643-3230.BCD-25-0264>
13. Lachowiec CA, Asimomitis G, Bernard E et al (2025) Multimodal and Data-Driven Assessment of Myeloid Neoplasms Refines Classification across Disease States. *Blood Cancer Discovery* 6:425–436. <https://doi.org/10.1158/2643-3230.BCD-25-0047>

14. Calvo X, Garcia-Gisbert N, Parraga I et al (2020) Oligomonocytic and overt chronic myelomonocytic leukemia show similar clinical, genomic, and immunophenotypic features. *Blood Adv* 4:5285–5296. <https://doi.org/10.1182/bloodadvances.2020002206>
15. Sango J, Carcamo S, Sirenko M et al (2024) RAS-mutant leukaemia stem cells drive clinical resistance to venetoclax. *Nature* 636:241–250. <https://doi.org/10.1038/s41586-024-08137-x>
16. Xu JJ, Chalk AM, Wall M et al (2022) Srsf2P95H/+ co-operates with loss of TET2 to promote myeloid bias and initiate a chronic myelomonocytic leukemia-like disease in mice. *Leukemia* 36:2883–2893. <https://doi.org/10.1038/s41375-022-01727-6>
17. Onida F, Kantarjian HM, Smith TL et al (2002) Prognostic factors and scoring systems in chronic myelomonocytic leukemia: a retrospective analysis of 213 patients. *Blood* 99:840–849. <https://doi.org/10.1182/blood.V99.3.840>
18. Such E, Germing U, Malcovati L et al (2013) Development and validation of a prognostic scoring system for patients with chronic myelomonocytic leukemia. *Blood* 121:3005–3015. <https://doi.org/10.1182/blood-2012-08-452938>
19. Elena C, Galli A, Such E et al (2016) Integrating clinical features and genetic lesions in the risk assessment of patients with chronic myelomonocytic leukemia. *Blood* 128:1408–1417. <https://doi.org/10.1182/blood-2016-05-714030>
20. Germing U, Strupp C, Aivado M, Gattermann N (2002) New prognostic parameters for chronic myelomonocytic leukemia? *Blood* 100:731–733. <https://doi.org/10.1182/blood-2002-01-0330>
21. Patnaik MM, Padron E, LaBorde RR et al (2013) Mayo prognostic model for WHO-defined chronic myelomonocytic leukemia: ASXL1 and spliceosome component mutations and outcomes. *Leukemia* 27:1504–1510. <https://doi.org/10.1038/leu.2013.88>
22. Patnaik MM, Itzykson R, Lasho TL et al (2014) ASXL1 and SETBP1 mutations and their prognostic contribution in chronic myelomonocytic leukemia: a two-center study of 466 patients. *Leukemia* 28:2206–2212. <https://doi.org/10.1038/leu.2014.125>
23. Lanino L, Hunter AM, Gagelmann N et al (2026) Molecular-Based Ecosystem to Improve Personalized Medicine in Chronic Myelomonocytic Leukemia. *J Clin Oncol* 44:1608–1623. <https://doi.org/10.1200/JCO-25-02116>
24. Saha A, Kakoty S, Patel K et al (2023) An Assessment of the Three Popular Prognostic Scoring Systems for Chronic Myelomonocytic Leukemia (CMML) in an Indian Context. *Indian J Med Pediatr Oncol* 44. <https://doi.org/10.1055/s-0043-1766130>
25. Srivastava VM, Nair SC, Sappani M et al (2023) Cytogenetic profile of 1791 adult acute myeloid leukemia in India. *Mol Cytogenet* 16:24. <https://doi.org/10.1186/s13039-023-00653-1>
26. Such E, Cervera J, Costa D et al (2011) Cytogenetic risk stratification in chronic myelomonocytic leukemia. *Haematologica* 96:375–383. <https://doi.org/10.3324/haematol.2010.030957>
27. Li MM, Datto M, Duncavage EJ et al (2017) Standards and Guidelines for the Interpretation and Reporting of Sequence Variants in Cancer: A Joint Consensus Recommendation of the Association

- for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists. *J Mol Diagn* 19:4–23. <https://doi.org/10.1016/j.jmoldx.2016.10.002>
28. Horak P, Griffith M, Danos AM et al (2022) Standards for the classification of pathogenicity of somatic variants in cancer (oncogenicity): Joint recommendations of Clinical Genome Resource (ClinGen), Cancer Genomics Consortium (CGC), and Variant Interpretation for Cancer Consortium (VICC). *Genet Sci* 24:986–998. <https://doi.org/10.1016/j.gim.2022.01.001>
29. Hunter JD (2007) Matplotlib: A 2D Graphics Environment. *Comput Sci Engg* 9:90–95. <https://doi.org/10.1109/MCSE.2007.55>
30. Waskom ML (2021) seaborn: statistical data visualization. *J Open Source Softw* 6:3021. <https://doi.org/10.21105/joss.03021>
31. Plotly Technologies Inc (2015) Collaborative Data Science. Accessed May 30, 2026. <https://plotly.com/>
32. Tang G, Zhang L, Fu B et al (2014) Cytogenetic Risk Stratification of 417 Patients with Chronic Myelomonocytic Leukemia from a Single Institution. *Am J Hematol* 89:813–818. <https://doi.org/10.1002/ajh.23751>
33. Tefferi A, Fathima S, Abdelmagid M et al (2025) BLAST: a globally applicable and molecularly versatile survival model for chronic myelomonocytic leukemia. *Blood* 146:874–886. <https://doi.org/10.1182/blood.2024027170>
34. Gurney M, Greipp PT, Gliem T et al (2023) TET2 somatic copy number alterations and allelic imbalances in chronic myelomonocytic leukemia. *Leuk Res* 134:107391. <https://doi.org/10.1016/j.leukres.2023.107391>
35. Garcia-Gisbert N, Arenillas L, Roman-Bravo D et al (2022) Multi-hit TET2 mutations as a differential molecular signature of oligomonocytic and overt chronic myelomonocytic leukemia. *Leukemia* 36:2922–2926. <https://doi.org/10.1038/s41375-022-01733-8>
36. Awada H, Nagata Y, Goyal A et al (2019) Invariant phenotype and molecular association of biallelic TET2 mutant myeloid neoplasia. *Blood Adv* 3:339–349. <https://doi.org/10.1182/bloodadvances.2018024216>
37. Itzykson R, Kosmider O, Renneville A et al (2013) Prognostic Score Including Gene Mutations in Chronic Myelomonocytic Leukemia. *J Clin Oncol* 31:2428–2436. <https://doi.org/10.1200/JCO.2012.47.3314>
38. Shomali W (2025) An adaptive survival prediction model in CMML. *Blood* 146:775–777. <https://doi.org/10.1182/blood.2025029619>
39. Loghavi S, Montalban-Bravo G (2026) Chronic myelomonocytic leukemia revisited: A comprehensive review with emphasis on the oligomonocytic subtype. *Hum Pathol* 106207. <https://doi.org/10.1016/j.humpath.2026.106207>

Figures

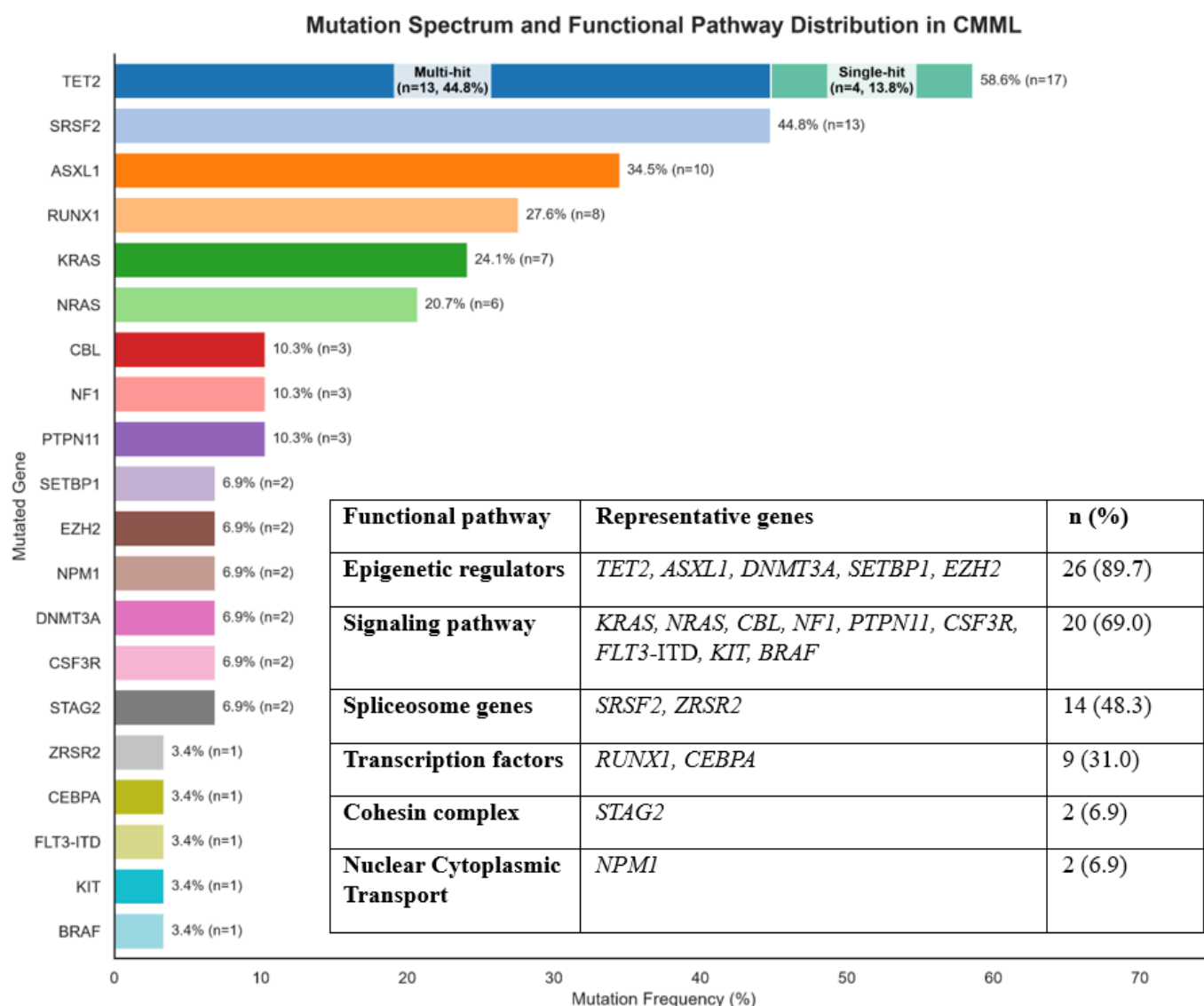


Figure 1

Molecular mutation spectrum and functional pathway distribution in CMML. Bars indicate the percentage (and number) of patients with mutations in each gene, as determined by targeted NGS. *TET2* alterations are shown as single-hit or putative multi-hit. The lower panel summarizes mutated genes according to the biological pathway

Comprehensive Oncoplot of CMML-1 versus CMML-2 (29 Cases)

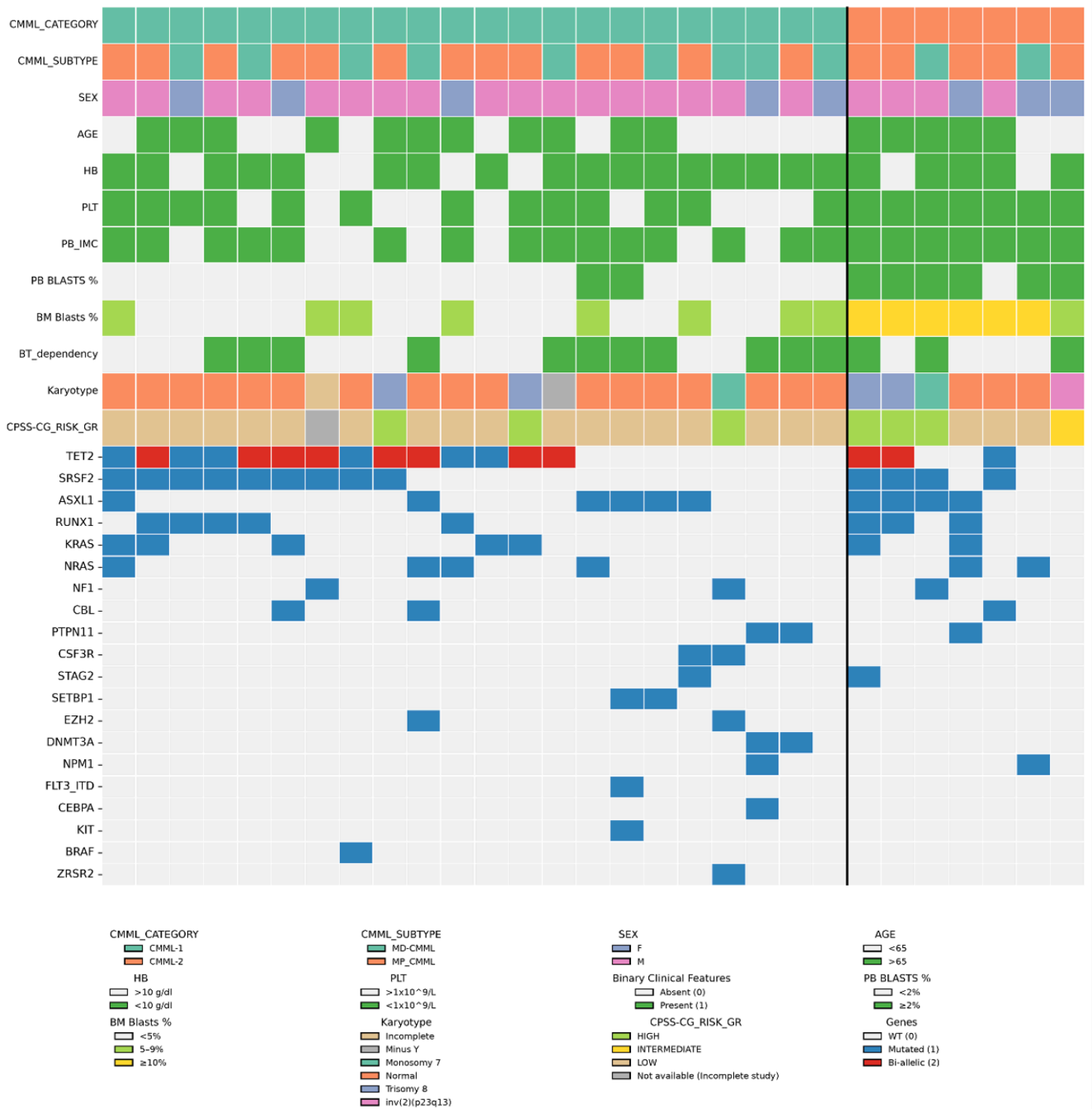


Figure 2

Integrated oncoplot of clinical, cytogenetic, and molecular features in CMML-1 versus CMML-2. Each column represents one patient, and each row represents a clinical feature, a cytogenetic category, a cytogenetic risk group, or a recurrent molecular alteration. Cases are grouped according to the WHO blast-defined category, with the vertical black line separating CMML-1 and CMML-2. Upper annotation

tracks show clinical variables, cytogenetic findings, and CPSS cytogenetic risk group, while lower tracks show recurrent gene mutations. *TET2* alterations are shown as single-hit or putative multi-hit.

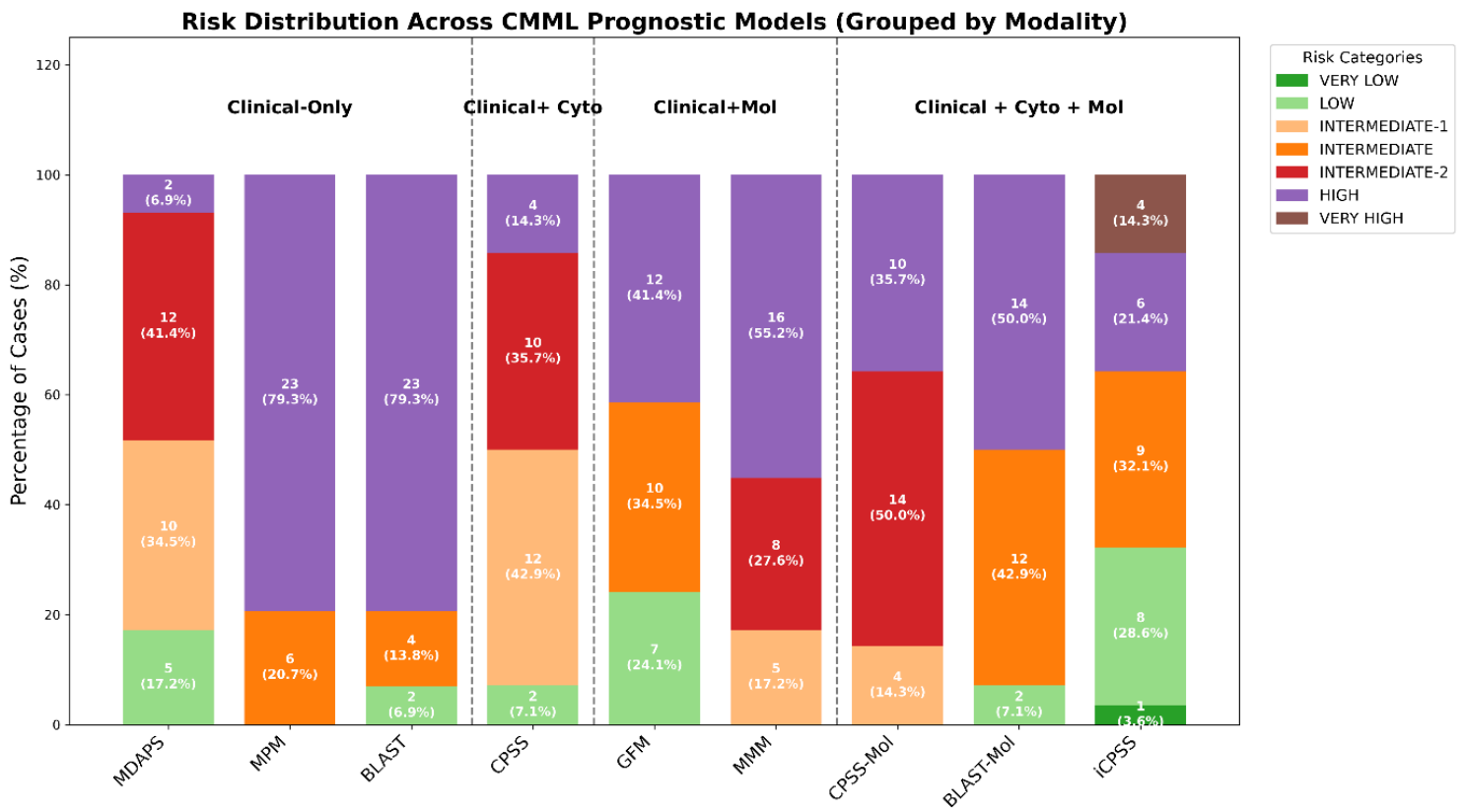


Figure 3

Risk distribution across CMML prognostic models grouped by modality. Stacked bar plot showing the percentage distribution of patients classified into each risk category by individual CMML prognostic scoring systems. Models are grouped into clinical-only, clinical plus cytogenetic (Clinical + Cyto), clinical plus molecular (Clinical + Mol), and integrated clinical–cytogenetic–molecular (Clinical + Cyto + Mol) scoring systems. Percentages are calculated using the number of cases with available data for each model; missing or non-evaluable cases were excluded from percentage calculation where applicable, and risk categories are color-coded from very low to very high risk.

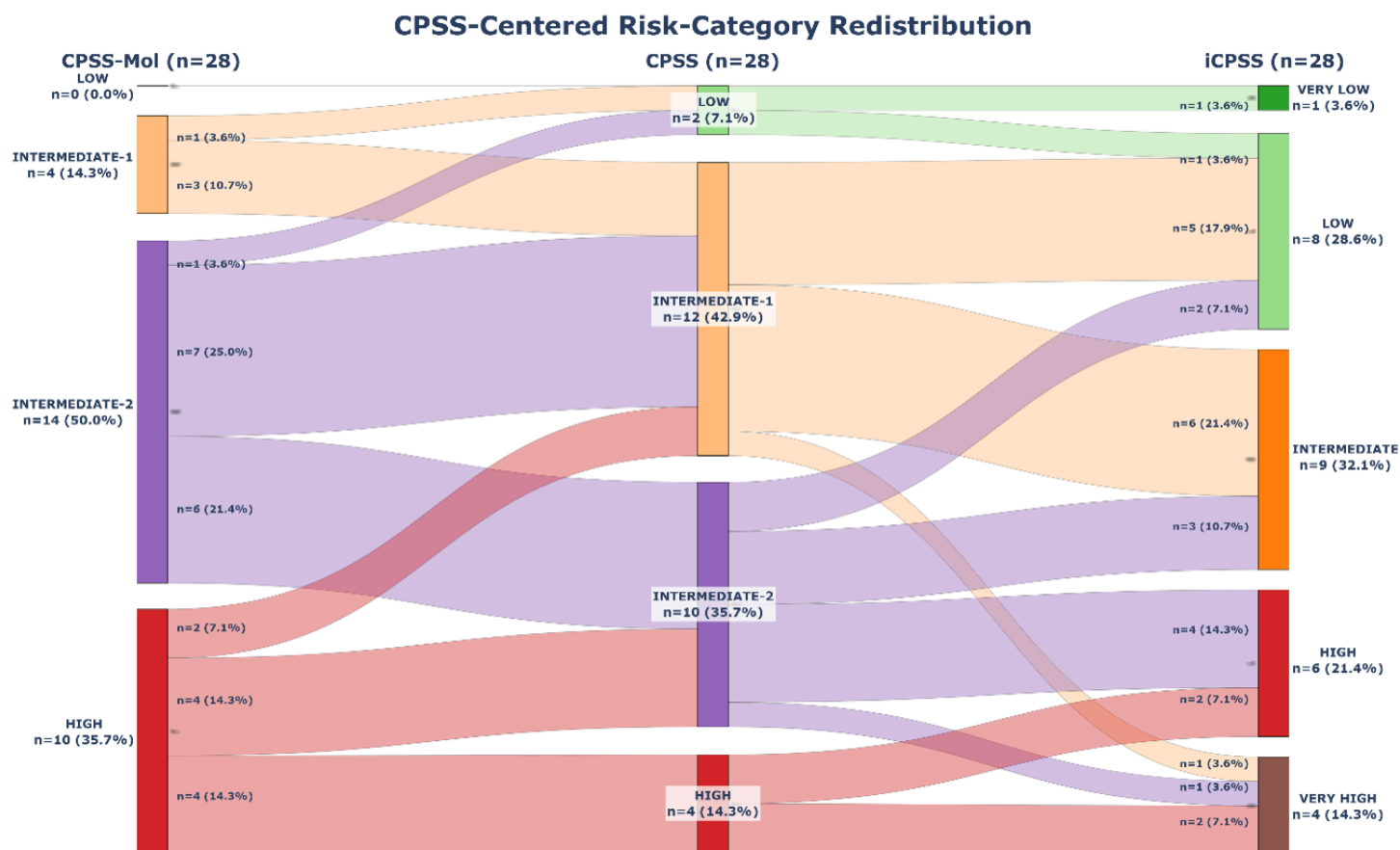


Figure 4

CPSS-centered risk-category redistribution. Alluvial plot showing baseline risk-class redistribution between CPSS and two integrated models, CPSS-Mol and iCPSS, among 28 evaluable CMML cases. CPSS is shown as the central reference model, with flows to CPSS-Mol on the left and iCPSS on the right. Each flow represents the number and percentage of patients moving between risk categories. Each flow represents the number of patients and the corresponding percentage of the 28 evaluable patients.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryInformationJHP.docx](#)